

HELSINGIN KAUPPAKORKEAKOULU  
Kansantaloustieteen laitos



TULOEROJEN JA TERVEYDEN VÄLINEN YHTEYS SUOMESSA

HELSINGIN  
KAUPPAKORKEAKOULUN  
KIRJASTO

10743

Kansantaloustieteen  
pro gradu -tutkielma  
Johanna Pekkala  
kevät 2008

Kansantaloustieteen

laitoksen johtajan päätöksellä 814 2008 hyväksytty

arvosanalla erinomainen, 90p.

Tarkastajat:

Prof. Pertti Haaparanta  
Prof. Pekka Ilmakunnas

11.3.2008

## TULOEROJEN JA TERVEYDEN VÄLINEN YHTEYS SUOMESSA

### Tavoitteet

Tutkielman tavoitteena oli selvittää, minkälainen yhteys maakuntien tuloeroilla oli suomalaisten pitkäaikaissairastavuuteen ja itse arvioituun yleiseen terveydentilaan vuosina 2003 ja 2004. Lisäksi tutkielmassa selvitettiin, oliko maakuntien tuloerojen yhteys yksilöiden terveyteen erilainen eri tulotasoilla olevien henkilöihin välillä. Tuloerojen ja terveyden välistä yhteyttä tarkasteltiin myös seutukunnittain. Tutkielmassa alueiden sisäisten tuloerojen mittaamisessa käytettiin kolmea eri eriarvoisuusmittaa.

### Tutkielman toteutustapa, aineisto ja menetelmät

Tutkielmassa testattiin empiirisen analyysin avulla tuloerojen ja terveyden välisestä yhteydestä esitettyä tulojen eriarvoisuuden hypoteesia, jonka käsittely sisältyi tutkielman teoreettiseen osioon. Tutkielman empiirisen osion poikkileikkausaineistona käytettiin Tilastokeskuksen vuosien 2003 ja 2004 tulonjakotilastoja, joihin yhdistettiin Euroopan Unionin tilastoviraston tulo- ja elinolotutkimusta varten tehdystä haastattelututkimuksesta saadut yksilöiden pitkäaikaissairastavuutta ja yleistä terveydentilaa koskevat tiedot. Aineisto analysoitiin probit-mallin ja järjestetyn probit -mallin avulla.

### Tulokset

Tutkielman tuloksista ilmeni, että maakunnan tuloeroilla ja pitkäaikaissairastavuudella ei ollut yhteyttä Suomessa vuosina 2003 ja 2004, kun tuloeroja mitattiin Gini-kertoimella. Tuloerojen ja pitkäaikaissairastavuuden välinen yhteys ei poikennut eri tulotasoilla olevien henkilöiden välillä tällöin. Tuloksista havaittiin kuitenkin, että maakunnan tuloeroilla ja itse arvioidulla yleisellä terveydentilalla oli käänteinen yhteys Suomessa vuonna 2003. Maakunnan sisäisten tuloerojen kasvaessa yleinen terveydentila heikkeni. Tuloerojen ja yleisen terveydentilan välinen yhteys ei kuitenkaan vaihdellut yksilöiden suhteellisen tulotason mukaan. Vuonna 2004 tuloerojen ja yleisen terveydentilan välillä ei esiintynyt yhteyttä. Kun tuloeroja tarkasteltiin seutukunnittain, havaittiin tuloerojen ja pitkäaikaissairastavuuden välillä positiivinen yhteys vuonna 2003. Tutkielman tulokset tukivat osin esitettyä hypoteesia tuloerojen kasvun terveyttä heikentävästä vaikutuksesta. Tutkielman empiiriset testit kuvaavat kuitenkin ensisijaisesti tuloerojen ja terveyden välistä korrelaatiota, eikä niiden perusteella voida tehdä luotettavia päätelmiä tuloerojen ja terveyden välisestä kausaalisuudesta. Tuloerojen ja terveyden välisestä yhteydestä tarvitaan lisää tietoa, jotta tutkimustuloksista voidaan tehdä päteviä johtopäätöksiä.

Avainsanat: tuloerot, eriarvoisuus, terveys, pitkäaikaissairastavuus, probit-malli, Suomi



# SISÄLLYSLUETTELO

1 JOHDANTO	5
1.1 Tutkielman taustaa	5
1.2 Lyhyt katsaus aikaisempiin tutkimuksiin	6
1.3 Tutkielman tavoitteet, tulokset ja sisältö	7
2 TULOEROT, ERIARVOISUUS JA TERVEYS SUOMESSA	10
2.1 Tuloerot ja eriarvoisuus	10
2.1.1 Käsitteitä	10
2.1.2 Tuloerojen mittaaminen	12
2.1.2.1 Lorenz-käytä ja Gini-kerroin	12
2.1.2.2 Sosiaalinen hyvinvointifunktio ja Atkinsonin indeksi	14
2.1.2.3 Yleinen entropiamitta	15
2.1.3 Köyhyyden mittaaminen	17
2.1.4 Eriarvoisuus ja köyhyys Suomessa 1990-luvulla ja 2000-luvun alussa	18
2.2 Väestön terveys	21
2.2.1 Terveiden määritelmästä ja mittaamisesta	22
2.2.2 Terveys ja sen väestöryhmittäiset erot Suomessa	23
3 TULOEROJEN JA TERVEYDEN VÄLISEN YHTEYDEN TEOREETTINEN TARKASTELU	26
3.1 Yleistä	26
3.2 Tulojen eriarvoisuuden hypoteesi	27
3.3 Vaikutusmekanismeja tuloerojen ja terveyden välillä	32
4 TULOEROT JA TERVEYS: KATSAUS EMPIIRISEEN KIRJALLISUUTEEN	36
4.1 Yleistä	36
4.2 Tuloerot ja terveys Suomessa	38
4.3 Tuloerot ja terveys Ruotsissa, Tanskassa ja Norjassa	41

5 AINEISTO, MUUTTUJAT JA TUTKIMUSMENETELMÄT	46
5.1 Aineisto	46
5.2 Tulokäsité ja ekvivalenssiskaalat	47
5.3 Muuttujien kuvaus	48
5.3.1 Selitettävät muuttujat	48
5.3.2 Aluepiirteitä kuvaavat selittävät muuttujat	53
5.3.3 Yksilön ja kotitalouden piirteitä kuvaavat selittävät muuttujat	55
5.4 Tutkimusmenetelmät	58
5.4.1 Probit-malli	59
5.4.2 Järjestetty probit -malli	62
6 TULOKSET	65
6.1 Tuloerojen yhteys pitkäaikaissairastavuuteen	65
6.2 Tuloerojen yhteys itse arvioituun yleiseen terveydentilaan	73
7 YHTEENVETO JA JOHTOPÄÄTÖKSET	80
LÄHTEET	86
LIITTEET	94



# 1 JOHDANTO

## 1.1 Tutkielman taustaa

Tuloeroissa on tapahtunut monia muutoksia viimeisimmän neljän vuosikymmenen aikana Suomessa. Tuloerot laskivat 1960- ja 1970-luvuilla, mutta kääntyivät lähihistorian ajan voimakkaimpaan nousuun 1990-luvun puolivälin tienoilla noin parinkymmenen vuoden tasaisen kehitysvaiheen jälkeen. Tuloerojen kasvu jatkui nopeana koko 1990-luvun jälkipuoliskon, mutta tasaantui 2000-luvun alkuvuosina. Tuloerot olivat Suomessa tällöin historiallisesti ennätyksellisen korkeat.<sup>1</sup> Tuloerojen ohella myös köyhyydessä tapahtui muutoksia 1990-luvun puolivälin jälkeen. Köyhien lukumäärän kasvun lisäksi myös itse köyhyys syveni Suomessa<sup>2</sup>.

Terveyserot ihmiskunnan huonoimmassa asemassa olevien köyhien ja hyvinvoivan eliitin välillä ovat tällä hetkellä historian suurimmat. Tämä terveyserojen kuvaus pätee myös Suomessa. Vaikka väestön terveys onkin kohentunut monilta osin Suomessa viime vuosikymmenien aikana, väestöryhmien välillä on merkittäviä terveyseroja useilla eri terveyden mittareilla tarkasteltuina. Esimerkiksi kuolleisuudessa sosioekonomisten ryhmien väliset erot ovat Suomessa melkein kaikkiin Länsi-Euroopan maihin verrattuna poikkeuksellisen suuret. Merkittävistä sosiaaliryhmien välisistä sairastavuuseroista johtuen erot terveisinä eletyissä elinvuosissa ovat elinajanodotettakin jyrkemmät eri sosiaaliryhmien välillä. Terveyserot väestöryhmien välillä eivät rajoitu pelkästään sosioekonomisen aseman mukaan havaittuihin eroihin, vaan niitä esiintyy myös esimerkiksi eri alueiden, sukupuolten ja siviilisäätöryhmien välillä.

Terveysongelmista aiheutuu yhteiskunnalle merkittäviä kustannuksia. Sairauksien aiheuttamat kustannukset syntyvät niin suorien kustannusten kuin menetettyjen työtulojen ja sitä kautta pienentyvien verotulojen kautta. Vanhasen II hallitus on lausunut hallitusohjelmassaan, että terveys on ”Suomen menestyksen kilpailutekijä, sillä kansanterveys on tärkeä taloudellisen, sosiaalisen ja inhimillisen kehityksen kulmakivi” (Pääministeri Matti Vanhasen II hallituksen ohjelma 2007, 74). Hallitus on todennut ohjelmassaan, että tulevana vuosikymmeninä väestörakenteen muuttuessa väestön hyvinvointi ja työkyky ovat ratkaisevan tärkeitä

---

<sup>1</sup> Riihelä ym. (2005, 1-2)

<sup>2</sup> Riihelä ym. (2007, 22)

yhteiskunnan toimivuuden kannalta (Pääministeri Matti Vanhasen II hallituksen ohjelma 2007, 74). Suomessa terveystalouden yhtenä tavoitteena onkin vähentää väestöryhmien eriarvoisuutta ja parantaa heikoimmassa asemassa olevien väestöryhmien hyvinvointia sekä pienentää sukupuolten ja eri koulutus- ja ammattiryhmien välisiä kuolleisuuseroja viidenneksellä vuoteen 2015 mennessä (Sosiaali- ja terveysministeriö 2001, 18).

Jotta sairauksia voidaan ehkäistä, terveyttä edistää ja terveyseroja kaventaa, on tunnettava ihmisen terveyteen vaikuttavia tekijöitä ja mekanismeja. Terveys on hyvin moniulotteinen käsite ja siihen vaikuttavat lukuisat eri tekijät. Yksilöllisillä tekijöillä on oma merkityksensä, mutta terveyteen vaikuttavat myös yksilön ulkoiset, ympäristöön liittyvät tekijät (Tremblay 2005). Esimerkiksi sosiaalisella ympäristöllä ja eriarvoisuudella on oma merkittävä vaikutuksensa (Tremblay 2005).

Käsillä olevan tutkielman päätavoitteena on selvittää, minkälainen yhteys tuloeroilla on yksilöiden terveyteen Suomessa. Lukuisilla sosioekonomisen aseman osoittimilla, kuten tuloilla, koulutuksella ja ammatilla on todettu olevan selvä ja jyrkkä yhteys terveyteen. Vaikka tulojen ja terveyden välillä havaitusta positiivisesta yhteydestä ollaan varsin yksimielisiä, tuloerojen ja terveyden välinen yhteys on kiistanalaisempi. Eri maissa julkaistut lukuisat tutkimukset eivät ole antaneet aiheesta yhdenmukaista kuvaa.

## 1.2 Lyhyt katsaus aikaisempiin tutkimuksiin

Tuloerojen ja terveyden välistä yhteyttä on tutkittu jo liki kolmen vuosikymmenen ajan. Aihetta ovat selvittäneet niin sosiologit, kansanterveystieteilijät kuin taloustieteilijätkin. Rodgers (1979) oli ensimmäisiä, joka tutki empiirisesti tuloerojen ja terveyden välistä yhteyttä eri maiden välillä ja havaitsi, että maan tuloeroilla oli yhteys elinajanodotteeseen ja lapsikuolleisuuteen. Tutkimuksessa suuret tuloerot olivat yhteydessä korkeaan kuolleisuuteen ja tasaisemman tulonjaon maissa elinajanodote oli tulojen suhteen eriarvoisempia maita korkeampi. Tämän jälkeen tutkimusten määrä on kasvanut huomattavasti. Wilkinson ja Pickett (2006) luettelevat kirjallisuuskatsauksessaan kaikkiaan 168 tuloerojen ja terveyden välistä yhteyttä selvittänyttä analyysia, joista hieman yli puolet tuki täysin esitettyä hypoteesia tuloerojen ja terveyden välisestä käänteisestä yhteydestä. Loput analyyseistä joko tukivat vain osin tai kumosivat sen.



Norjaa lukuun ottamatta useimmissa pohjoismaalaisissa tutkimuksissa tuloerojen ja terveyden välillä ei ole havaittu yhteyttä (ks. Osler ym. 2002, Blomgren ym. 2004, Gerdtham & Johannesson 2004, Martikainen ym. 2004, Henriksson ym. 2006, Böckerman ym. 2007). Pohjoismaat ovat monessa suhteessa samankaltaisia: esimerkiksi tuloerot ovat kansainvälisesti vertaillen varsin alhaiset. Esitetyn väitteen mukaan tuloerojen ja terveyden välillä olisi yhteys ainoastaan maissa, joissa eriarvoisuus on suurta (ks. Subramanian & Kawachi 2004). Muutamit Suomessa julkaistut tuloerojen ja terveyden välistä yhteyttä selvittäneet tutkimukset ovat olleet tämän väitteen mukaisia (ks. Blomgren ym. 2004, Martikainen ym. 2004, Böckerman ym. 2007). Böckerman ym. (2007) havaitsivat tosin, että naisten kohdalla tuloeroilla oli käänteinen yhteys itse arvioituun hyvään fyysiseen terveydentilaan ja ei-työkyvyttömyyseläkkeellä oloon vuosina 1993–2005. Koko väestötasolla ei vastaavanlaista yhteyttä ole kuitenkaan Suomessa todettu. Koska tuloerojen ja terveyden välinen yhteys on osoittautunut kansainvälisissä tutkimuksissa varsin kiistanalaiseksi, eikä aihetta ole tutkittu Suomessa kovinkaan paljon tuloerojen viime vuosien varsin voimakkaasta kasvusta ja väestöryhmien välisistä merkittävistä terveyseroista huolimatta, tarvitaan runsaasti lisätietoa ennen kuin lopullisia johtopäätöksiä taloudellisen eriarvoisuuden ja terveyden välisestä yhteydestä Suomessa voidaan tehdä.

### 1.3 Tutkielman tavoitteet, tulokset ja sisältö

Käsillä olevan tutkielman keskeisimpänä tavoitteena on selvittää, minkälainen tuloerojen ja terveyden välinen yhteys on Suomessa. Tutkimusongelma voidaan tiivistää seuraaviin kysymyksiin ja kirjallisuudessa esitettyihin hypoteeseihin:

1. Minkälainen yhteys maakuntien tuloerojen ja yksilöiden terveyden välillä oli Suomessa vuosina 2003 ja 2004? Onko tuloerojen ja terveyden välinen yhteys erilainen, jos tuloerot lasketaan seutukunnittain? Entä onko yhteys erilainen, mikäli tuloerojen mittaamisessa käytetään eri tuloeromittoja tai erilaisia terveyttä kuvaavia indikaattoreita?

*Hypoteesi 1.* Tuloerot ovat haitallisia kaikkien yksilöiden terveydelle.

2. Oliko maakuntien tuloerojen yhteys yksilöiden terveyteen erilainen eri tulotasoilla olevien henkilöihin välillä vuosina 2003 ja 2004? Muuttuuko tutkimustulos, jos estimoinnissa käytetään erilaisia terveyttä kuvaavia indikaattoreita?

*Hypoteesi 2.* Tuloerot ovat haitallisimpia tulojen suhteen huono-osaisten terveydelle.

Käsillä olevassa tutkielmassa testataan empiirisesti tulojen eriarvoisuuden hypoteesia (engl. income inequality hypothesis). Edellä esitetystä hypoteesista 1 on käytetty kirjallisuudessa vapaasti suomennettuna nimitystä tulojen eriarvoisuuden hypoteesin vahva versio (engl. the strong version of the income inequality hypothesis / the strong income inequality hypothesis). Hypoteesia 2 on vapaasti suomennettuna nimitetty tulojen eriarvoisuuden hypoteesin heikoksi versioksi (engl. the weak version of the income inequality hypothesis / the weak income inequality hypothesis).

Lisäksi tutkielmassa on tavoitteena selvittää, minkälainen yhteys maakunnittain lasketulla köyhyydellä on yksilöiden terveyteen. Tätä kysymystä lähestytään tuloerojen näkökulmasta siten, että selvitetään, muuttuuko havaittu tuloerojen ja terveyden välinen yhteys, mikäli samanaikaisesti otetaan huomioon myös alueen köyhyys.

Aineistona tutkielmassa käytetään Tilastokeskuksen vuosien 2003 ja 2004 tulonjakotilastoja, joihin yhdistetään Euroopan Unionin tilastoviraston tulo- ja elinolotutkimusta varten tehdystä haastattelututkimuksesta SILC-aineistosta saadut yksilöiden pitkäaikaissairastavuutta ja yleistä terveydentilaa koskevat tiedot. Käsillä oleva tutkielma on ensimmäinen Suomessa tehty tutkimus, jossa hyödynnetään SILC-aineistosta saatavia tietoja tuloerojen ja terveyden välisen yhteyden selvittämisessä. Vastaavia tutkimuksia ei aineistosta ole aiemmin tehty.

Tutkielman tuloksista ilmenee, että maakunnan tuloeroilla ja pitkäaikaissairastavuudella ei ollut yhteyttä Suomessa vuosina 2003 ja 2004, kun tuloeroja mitattiin Gini-kertoimella. Vuonna 2003 yhteys oli positiivinen, kun Gini-kerroin korvattiin Theilin indeksillä. Myös seutukunnan tuloeroille ja pitkäaikaissairastavuudelle saatiin tilastollisesti merkitsevä positiivinen yhteys vuonna 2003. Tutkielman toisen keskeisen tuloksen mukaan maakunnan tuloeroilla ja yleisellä terveydentilalla oli negatiivinen yhteys Suomessa vuonna 2003 riippumatta valitusta eriarvoisuusmitasta. Tuloerot olivat haitallisia yleiselle terveydentilalle. Tutkielman tulokset osoittivat kuitenkin, että tuloerojen ja yleisen terveydentilan välinen yhteys ei vaihdellut yksilöiden suhteellisen tulotason mukaan. Seutukunnan tuloerojen ja itse



arviodun yleisen terveydentilan välillä yhteyttä ei esiintynyt vuosina 2003 ja 2004. Tulokset säilyivät pääpiirteissään muuttumattomina, vaikka samanaikaisesti huomioitiin alueen köyhyys. Tutkielman tulokset tukivat osin tulojen eriarvoisuuden hypoteesin vahvaa versiota, mutta hypoteesin heikkoa versiota tulokset eivät vahvistaneet.

Tutkielma etenee seuraavasti. Luku 2 jakautuu kahteen pääteemaan. Sen ensimmäisessä alaluvussa käsitellään tuloeroja ja eriarvoisuutta. Tarkastelu aloitetaan käsitteiden määrittelyllä, minkä jälkeen selvitetään tuloerojen ja köyhyyden mittaamista. Ensimmäinen alaluku päättyy kuvaukseen tuloeroista ja köyhyydestä Suomessa 1990-luvun alusta vuoteen 2004. Luvun 2 toinen alaluku käsittelee väestön terveyttä Suomessa. Luvussa perehdytään terveyden määritelmään ja mittaamiseen, minkä jälkeen luodaan katsaus väestön terveyteen ja terveyden väestöryhmittäisiin eroihin Suomessa.

Kolmannen luvun alussa luodaan katsaus kirjallisuudessa esitettyihin hypoteeseihin tulojen, köyhyyden ja tuloerojen yhteydestä terveyteen. Tämän jälkeen tarkastelussa keskitytään tulojen eriarvoisuuden hypoteesiin. Luvussa 3.2 esitetään teoreettinen malli tuloerojen ja terveyden välisestä yhteydestä ja perehdytään tarkemmin tulojen eriarvoisuuden hypoteesin kahteen eri versioon. Luvun 3 lopussa käsitellään erilaisia, kirjallisuudessa esitettyjä vaikutusmekanismeja tuloerojen ja terveyden välillä. Luvussa 4 luodaan katsaus tuloerojen ja terveyden välistä yhteyttä selvittäneeseen empiiriseen kirjallisuuteen.

Luku 5 sisältää kuvauksen tutkielman empiirisessä osiossa käytettävästä aineistosta, tulokäsitteestä, ekvivalenssiskaaloista sekä muuttujista. Luvun 5 lopuksi esitellään estimoinnissa käytettävät probit-malli ja järjestetty probit-malli sekä selvitetään estimoitavien parametrien tulkintaa. Luvussa 6 esitetään mallien estimointitulokset ja tulosten tulkinta. Tutkielman viimeinen luku sisältää yhteenvedon ja johtopäätökset.

## 2 TULOEROT, ERIARVOISUUS JA TERVEYS SUOMESSA

### 2.1 Tuloerot ja eriarvoisuus

#### 2.1.1 Käsitteitä

Kotitalouden tulot voidaan määritellä usealla eri tavalla. Seuraavassa tarkastellaan tutkielman kannalta keskeisiä tulokäsitteitä. Tuotannontekijätulot muodostuvat palkka-, yrittäjä- ja omaisuustuloista. Näistä tulolajeista on käytetty myös nimitystä markkinatulot. Bruttotulot saadaan, kun tuotannontekijätuloihin lisätään saadut tulonsiirrot kuten esimerkiksi sosiaaliturvaetuudet ja sosiaaliavustukset. Kun bruttotuloista vähennetään maksetut tulonsiirrot, päästään käytettävissä oleviin tuloihin eli nettotuloihin. Rahatulot saadaan puolestaan, kun bruttotuloista vähennetään laskennalliset tuloerät. Näitä eriä ovat esimerkiksi omassa käytössä olevasta omistusasunnosta saadut laskennalliset tulot.<sup>3</sup>

Henkilökohtaista tulonjakoa ja köyhyyttä selvittämissä tutkimuksissa havaintoyksikkönä tarkastellaan yleensä kotitalouden käytettävissä olevia tuloja eli nettotuloja. Yksilöt käyttävät myös erilaisia ilmaisia tai subventoituja julkisia palveluita, kuten koulutus-, sosiaali- ja terveyspalveluja. Näille palveluille on mahdollista laskea arvo, jota voidaan tarkastella yhtenä tuloeränä. Kokonaistulot saadaan, kun käytettävissä oleviin tuloihin lisätään asumisen korkotukien lisäksi käytettyjen palveluiden arvo, josta vähennetään palveluista maksetut käyttäjämaksut (Loikkanen ym. 1997, 6). Kokonaistuloja ei tarkastella tässä tutkielmassa, sillä kyseessä oleva tulokäsite ei sisälly tutkielman empiirisessä osiossa käytettävään tulonjakotilaston aineistoon.

Koska kotitaloudet ansaitsevat eri määrän tuloja, tulot jakautuvat epätasaisesti kotitalouksien kesken. Tuloeroja tutkittaessa on tulojakauma tutkimuksen lähtökohta. Tulojakaumassa esiintyviä tuloeroja voidaan tutkia kahdella tavalla. Voidaan tarkastella joko tulojen suhteellista vaihtelua tai tutkia tulojen absoluuttisia eroja.

---

<sup>3</sup> Aineistona tutkielmassa käytetään Tilastokeskuksen vuosien 2003 ja 2004 tulonjakotilastoja, joiden keskeisiä tulokäsitteitä luettiin edellä.



Tuloeroja mitataan eriarvoisuusmitoilla, jotka kuvaavat jakauman tuloerojen suuruutta. Tulojen eriarvoisuutta mitataan tutkielmassa kolmella suhteellisella eriarvoisuusmitalla: Gini-kertoimella, Atkinsonin indeksillä ja yleisellä entropiamitalla. Eriarvoisuusmittojen avulla tulojakaumille voidaan antaa tunnusluku ja järjestää siten jakaumat niiden sisältämän eriarvoisuuden suuruuden mielessä järjestykseen. Tässä tutkielmassa lähtökohtaoletuksena on, että tulot ovat taloudellinen resurssi, jonka suhteen eriarvoisuutta tarkastellaan.

Tulkintaa eriarvoisuudesta voidaan pitää tuloerojen suuruuteen liittyvänä normatiivisena pohdintana. Taloustieteessä eriarvoisuudella tarkoitetaan usein mahdollisuuksien tai toteutuneen elintason eriarvoisuutta. Toisaalta eriarvoisuuteen ei välttämättä liitetä minkäänlaisia eettisiä näkökulmia, vaan sillä tarkoitetaan ainoastaan havaittuja tuloeroja.

Eriarvoisuutta tutkittaessa ollaan kiinnostuneita koko tulojakaumasta ja siinä esiintyvistä tuloeroista. Köyhyyttä selvitettäessä tarkastelun kohteena on tulojakauman tietyn rajan alapuolella oleva osa. Köyhyys on näin ollen eriarvoisuutta rajatumpi tutkimuskohde. Tuloerojen tavoin myös köyhyyttä voidaan tarkastella joko suhteellisesta tai absoluuttisesta näkökulmasta. Suhteellisesta köyhyydestä puhuttaessa verrataan pienituloisen väestöosan tuloja muun väestön tuloihin. Tällöin tuloja verrataan köyhyysrajaan, joka on määritelty tietyssä osuutena esimerkiksi tulojakauman keskimääräisestä tulosta (Foster 1998, 336). Absoluuttinen köyhyysraja on puolestaan määrätty kiinteä tulotaso, joka ei vaihtelee tulojakauman mediaani- tai keskiarvotulon muuttuessa (ma. 336).

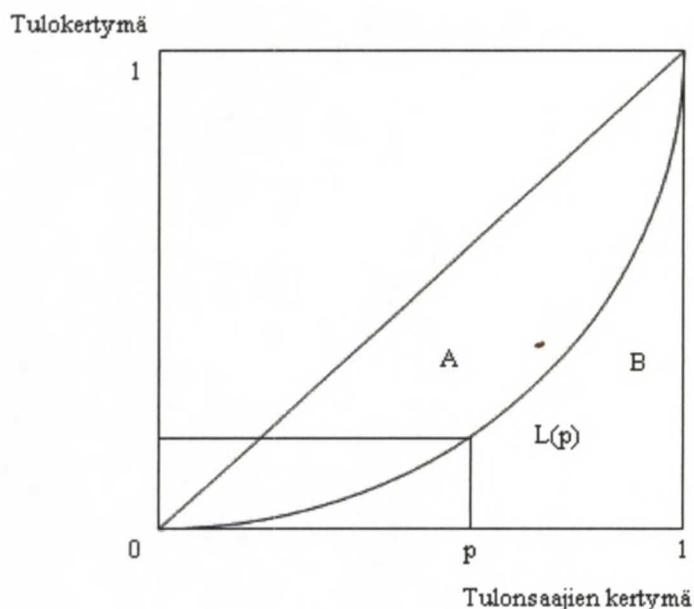
## 2.1.2 Tulorojen mittaaminen

Seuraavassa käsitellään kolmea yleisesti käytettyä suhteellista eriarvoisuusmittaa. Ensimmäisenä tarkastellaan Lorenz-käyrää ja siihen perustuvaa Gini-kerrointa, minkä jälkeen selvitetään sosiaalista hyvinvointifunktiota ja Atkinsonin eriarvoisuusindeksiä. Viimeisenä käsitellään yleistä entropiamittaa ja sen erityistapausta Theilin indeksiä.

### 2.1.2.1 Lorenz-käyrä ja Gini-kerroin<sup>4</sup>

Tulojen jakautumista voidaan tarkastella Lorenz-käyrän avulla. Lorenz-käyrä  $L(p)$  havainnollistetaan kuvaajana neliössä, jonka akselien pituus on yksi (ks. kuvio 1). Käyrä saadaan järjestämällä tulonsaajat tulojen mukaan nousevaan järjestykseen ja selvittämällä, minkä osuuden vähiten ansaitseva tulonsaajien ryhmä saa. Kuviossa 1 vaaka-akseli kuvaa tulonsaajien kertymää ja pystyakseli kyseessä olevien tulonsaajien tulokertymää. Lorenz-käyrä  $L(p)$  kuvaa sitä osuutta kokonaistuloista, jonka köyhimmät 100p prosenttia ihmisistä saa, kun  $p \in (0, 1)$ .

Kuvio 1. Lorenz-käyrä  $L(p)$



<sup>4</sup> Lorenz-käyrän ja Gini-kertoimen teoriasta Lambert (1993, 30–47).

Kuviossa 1 diagonaalisuora kuvaa tilannetta, jossa Lorenz-käyrä on yhtenevä  $45^\circ$ -suoran kanssa. Tällöin tulonsaajien kumulatiivinen osuus on yhtä suuri kuin kyseessä olevien tulonsaajien tulojen kertymä. Toisin sanoen kaikkien tulot ovat tällöin yhtä suuret. Koska tulot ovat jakautuneet tasaisesti kaikkien tulonsaajien kesken, ei tuloeroja ole. Tuloerojen ollessa suurimmillaan Lorenz-käyrän poikkeama  $45^\circ$ -suorasta on maksimissa. Tällöin Lorenz-käyrä kulkee ensin vaaka-akselilla ja sen jälkeen pystyakselia pitkin kohtisuoraan ylöspäin. Tällöin suurituloisin saa kaikki tulot. Mitä suuremmat tuloerot ovat, sitä kauempana Lorenz-käyrä sijaitsee tasaista tulonjakoa kuvaavasta  $45^\circ$ -suorasta.

Lorenz-käyrän avulla ei voida tehdä päätelmiä tulojakaumien paremmuudesta hyvinvoinnin osalta. Kahden tulojakauman Lorenz-käyrät voivat olla yhtenevät, vaikka toisen jakauman tulot olisivat absoluuttisesti kaksi kertaa suuremmat kuin toisen. Lorenz-käyrän asemaan eivät vaikuta kaikkien jakauman tulojen yhtä suuret suhteelliset muutokset. Tulojen suhteelliset erot voivat olla kahdessa tulojakaumassa samat, vaikka rikkaiden ja köyhien absoluuttisten tuloerojen suuruus poikkeaisi jakaumien kesken.

Tulojen eriarvoisuuden kuvaamisessa kenties yleisimmin käytetty mitta on Gini-kerroin. Gini-kerroin kuvaa sen alueen kokoa, jonka Lorenz-käyrä poikkeaa tasaista tulonjakoa kuvaavasta  $45^\circ$ -suorasta. Gini-kerroin lasketaan jakamalla Lorenz-käyrän ja tasaista tulonjakoa kuvaavan  $45^\circ$ -suoran väliin jäävä alue (alue A kuviossa 1) koko sillä alueella, joka jää  $45^\circ$ -suoran alapuolelle (alue A+B kuviossa 1). Kuvion 1 merkintöjä käyttäen Gini-kerroin on:

$$G = \frac{A}{A+B} = 2A = 2\left(\frac{1}{2} - B\right) = 1 - 2B. \quad (1)$$

Mikäli Lorenz-käyrän funktio tunnetaan, Gini-kerroin voidaan laskea pinta-alaintegraalin avulla. Yhtälön 1 perusteella Gini-kerroin saadaan, kun ykkösestä vähennetään kaksi kertaa Lorenz-käyrän alapuolella oleva alue B. Pinta-ala B saadaan laskemalla määrätty integraali Lorenz-käyrän funktiosta  $L(p)$  välillä  $[0, 1]$ . Tällöin Gini-kerroin on:

$$G = 1 - 2 \int_0^1 L(p) dp. \quad (2)$$

Gini-kerroin on mahdollista esittää myös tulojen ja tulojärjestyksen kovarianssina:



$$G = (2/\mu) \text{cov}\{y, F(y)\}, \quad (3)$$

jossa  $y$  on tulot,  $\mu$  tulojen aritmeettinen keskiarvo ja  $F(y)$  tulon  $y$  kertymäfunktio.

Gini-kerroin saa arvoja 0 ja 1 välillä siten, että tuloerot kasvavat Gini-kertoimen lähestyessä arvoa yksi. Lorenz-käyrän poikkeama tasaista tulonjakoa kuvaavasta 45°-suorasta on suurimmillaan Gini-kertoimen ollessa 1. Tällöin suurituloisin saa kaikki tulot ja tuloerot ovat suurimmillaan. Kun Gini-kerroin saa arvon 0, Lorenz-käyrä on yhtenevä 45°-suoran kanssa. Tällöin kaikkien tulot ovat yhtä suuret eikä tuloeroja ole. Gini-kerroin reagoi herkimmin tulojakauman mooditulon läheisyydessä oleviin tuloeroihin (Jenkins 1999, 8).

#### 2.1.2.2 Sosiaalinen hyvinvointifunktio ja Atkinsonin indeksi<sup>5</sup>

Atkinsonin (1970) eriarvoisuusindeksi kuuluu eettisiin tuloeromittareihin. Kyseessä olevat tuloeromittarit pyrkivät yhdistämään jakauman eriarvoisuuden tuloeroista aiheutuvaan hyvinvoinnin tappioon. Atkinsonin indeksi perustuu sosiaaliseen hyvinvointifunktioon, jonka mukaan yhteiskunnan sosiaalinen hyvinvointi on yksilöiden sosiaalisen hyödyn summa.

Oletetaan, että yksilön sosiaalinen hyötyfunktio on kaikille yksilöille samanlainen. Yksilön sosiaalinen hyöty määräytyy tulojen mukaan siten, että hyvinvointi kasvaa tulojen noustessa. Koko yhteiskunnan sosiaalinen hyvinvointi  $W$  saadaan:

$$W = \sum u(y_i), \quad (4)$$

jossa  $u(y_i)$  on yksilöiden sosiaalinen hyötyfunktio ja  $y_i$  yksilön  $i$  tulot.

Tulojen jakautuessa tasan jokaisen yksilön sosiaalinen hyöty on  $u(\bar{y})$ , jossa  $\bar{y}$  on keskimääräinen tulotaso. Vastaavasti sitä tuloa, joka tasan jakautuessaan tuottaisi yhteiskunnalle saman sosiaalisen hyvinvoinnin kuin alkuperäinen eriarvoinen tulojakauma, nimitetään saman hyvinvoinnin tason tuottavaksi tasaisesti jaetuksi tuloksi (engl. equally

<sup>5</sup> Sosiaalisesta hyvinvointifunktiosta ja Atkinsonin indeksistä Cowell (1995, 36–47).

distributed equivalent income)  $y_e$ <sup>6</sup>. Atkinsonin indeksi määritellään keskiarvotulon ja saman hyvinvoinnin tason tuottavan tasaisesti jaetun tulon välisenä suhteellisenä erona:

$$I = 1 - \frac{y_e}{\bar{y}}. \quad (5)$$

Atkinsonin indeksin arvot vaihtelevat 0 ja 1 välillä siten, että tulojen eriarvoisuuden kasvaessa indeksi lähestyy arvoa 1 (Cowell 1995, 142). Keskimääräisen tulon ja saman hyvinvoinnin tason tuottavan tasaisesti jaetun tulon suhteellinen ero yhtälössä 5 kuvaa sitä tulon määrää, joka voitaisiin hävittää ilman sosiaalista hyvinvointitappiota, mikäli loput jäljelle jääneistä tuloista jaettaisiin tasan. Kun yhteiskunnan eriarvoisuuden kaihtamisen aste nousee (parametri  $\epsilon$  kasvaa, ks. alaviite 6), saman hyvinvoinnin tason tuottava tasaisesti jaettu tulo pienenee ja Atkinsonin indeksi kasvaa (Duclos & Araar 2006, 64–65). Mitä suurempi yhteiskunnan eriarvoisuuden kaihtamisen aste on, sitä herkemmin Atkinsonin indeksi reagoi tuloeroihin tulojakauman alaosassa (Jenkins 1999, 8).

### 2.1.2.3 Yleinen entropiamitta<sup>7</sup>

Informaatioteorian näkökulmasta entropialla viitataan tietyn tapahtumajoukon ( $i = 1, 2, 3, \dots, n$ ) informaattiosisältöön. Tapahtumaan  $i$  liittyvää todennäköisyyttä kuvataan muuttujalla  $p_i$ . Funktiolla  $h(p_i)$  viitataan sen informaation arvoon, joka liittyy kyseessä olevan tapahtuman todennäköisyyteen. Funktion  $h(p_i)$  arvo kasvaa tapahtuman todennäköisyyden  $p_i$  pienentyessä. Koko tapahtumajoukon keskimääräinen informaattiosisältö eli entropia lasketaan yksittäisten tapahtumien informaatioarvon tapahtumien todennäköisyyksillä painotettuna summana:

$$\text{entropia} = \sum_{i=1}^n p_i h(p_i). \quad (6)$$

<sup>6</sup>  $y_e = [\sum f_i (y_i)^{1-\epsilon}]^{1/(1-\epsilon)}$ , kun  $\epsilon \geq 0$ ,  $\epsilon \neq 1$  ( $i = 1, \dots, n$ ).  $y_e = \sum f_i \log(y_i)$ , kun  $\epsilon = 1$ . (Jenkins 1999, 8–9) Tutkielman otos korotetaan perusjoukon tasolle kotitalouskohtaisten painokertoimien avulla. Tällöin  $f_i = h_i w_i / N$ , jossa  $N = \sum h_i w_i$  (Jenkins 1999, 8). Muuttuja  $h_i$  on kotitalouden  $i$  jäsenmäärä ja  $w_i$  kotitalouskohtainen korotuskerroin.

<sup>7</sup> Yleisestä entropiamitasta Cowell (1995, 47–54).

Funktio  $h$  esitetään muodossa  $h = -\log(p_i)$ , sillä vain tätä muotoa käyttämällä riippumattomien tapahtumien  $i$  ja  $j$  todennäköisyyttä  $p_i p_j$  voidaan tarkastella muodossa:

$$h(p_i p_j) = h(p_i) + h(p_j). \quad (7)$$

Sovellettaessa informaatioteorian entropia-käsitettä eriarvoisuuden tutkimiseen korvataan tapahtuman todennäköisyys  $p_i$  henkilön  $i$  tulo-osuudella  $s_i$ :

$$s_i = \frac{y_i}{ny}, \quad (8)$$

jossa  $y_i$  on henkilön  $i$  tulot,  $n$  henkilöiden kokonaismäärä ja  $\bar{y}$  tulojen keskiarvo siten, että

$$\sum_{i=1}^n s_i = 1. \quad (9)$$

Vähennettäessä tulojakauman todettu entropia entropian maksimiarvosta (kun kaikkien henkilöiden tulo-osuus on yhtä suuri eli  $s_i = \bar{y}/(n\bar{y}) = 1/n$ ), saadaan eriarvoisuusindeksi:

$$\begin{aligned} T &= \sum_{i=1}^n \frac{1}{n} h\left(\frac{1}{n}\right) - \sum_{i=1}^n s_i h(s_i) \\ &= \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i}{y}\right) \log\left(\frac{y_i}{y}\right). \end{aligned} \quad (10)$$

Kaavassa 10 esitetty Theilin (1967) indeksi  $T$  on erityistapaus yleisestä entropiamitasta<sup>8</sup>, joka voidaan esittää muodossa:

$$GE(\theta) = \frac{1}{\theta(1-\theta)} \left[ \left[ \frac{1}{n} \sum_{i=1}^n (y_i / \bar{y})^\theta \right] - 1 \right], \quad (11)$$

<sup>8</sup> Yleinen entropiamitta ja sen erityistapaus Theilin indeksi esitetään tässä painottamattomina. Tutkielman empiirisessä osiossa Theilin indeksin laskemisessa käytetään painotettua aineistoa. Tällöin indeksit korotetaan koko perusjoukon tasolla siten, että esitetyssä kaavassa 10  $(1/n)$  korvataan  $f_i = h_i w_i / N$ , jossa  $N = \sum h_i w_i$  ja  $i = 1, \dots, n$  (Jenkins 1999, 8).



missä  $\theta \neq 0,1$ . Parametrin  $\theta$  arvolla 1 saadaan Theilin indeksi  $T$ . Yleisen entropiamitan arvot vaihtelevat välillä  $[0, \infty[$  siten, että tulojen jakautuessa tasan indeksi saa arvon 0 (Cowell 1995, 142).

Yleisen entropiamitan herkkyys tulojakauman eri osissa esiintyviin tuloeroihin riippuu parametrin  $\theta$  arvosta. Mitä suurempia arvoja parametri  $\theta$  saa, sitä herkemmin mitta reagoi tulojakauman yläosissa oleviin tuloeroihin ja päinvastoin (Jenkins 1999, 8). Kun  $\theta = 1$  (Theilin indeksi), tulopoikkeamat saavat yhtä suuren painon koko jakaumassa<sup>9</sup>.

### 2.1.3 Köyhyyden mittaaminen

Seuraavassa tarkastellaan köyhyyden laskemisessa käytettyjä menetelmiä<sup>10</sup>. Tarkastelun kohteena on tulojakauman tietyn rajan alapuolella oleva osa. Suhteellisen köyhyyden määrittelyssä yleisesti käytetty periaate on tulojen vertaaminen köyhyysrajaan, joka määritellään tietyssä prosenttiosuutena keskimääräisestä kansallisesta tulotasosta<sup>11</sup>. Köyhiksi luokitellaan tällöin ne yksilöt tai kotitaloudet, joiden tulot alittavat köyhyysrajan. Köyhyysraja voidaan laskea joko tulojen keskiarvosta tai tulojen mediaanista. Myös köyhyysrajan määrittelyssä käytetyt prosenttiosuudet vaihtelevat.

Köyhyyden mittaamisessa yleisesti käytetty menetelmä on päälukumitta (engl. head count ratio). Se ilmoittaa köyhyysrajan alapuolella olevien tulonsaajien osuuden kaikista tulonsaajista eli köyhien väestöosuuden:

$$H = \frac{1}{n} \sum_{i=1}^n 1(y_i \leq z), \quad (12)$$

jossa  $1(\cdot)$  saa arvon 1, jos henkilö on köyhyysrajalla tai sen alapuolella ja 0, mikäli tilanne on päinvastoin. Atkinson (1998, 49) esittää, että päälukumitta on mielekäs köyhyysmittari, mikäli köyhyysrajaa tarkastellaan perusoikeuksiin kuuluvana vähimmäistulona. Tällöin päälukumitta ilmoittaa niiden henkilöiden osuuden, joiden kohdalla vähimmäistulovaatimus

<sup>9</sup> Kun parametri  $\theta$  saa arvon 0 (logaritminen keskipoikkeama), yleistetty entropiamitta antaa suhteellisesti enemmän painoa tulopoikkeamille jakauman alaosassa. Kun  $\theta$  on 2 (variaatiokertoimen neliön puolikas), indeksi painottaa enemmän tulopoikkeamia jakauman yläosassa. (ks. Riihelä & Sullström 2001, 22)

<sup>10</sup> Köyhyyden mittaamisesta esim. Atkinson (1998, 48–52), Riihelä ym. (2007, 2–6), Suoniemi (2005, 25).

<sup>11</sup> Tulojen lisäksi köyhyys voidaan määritellä myös kulutusmenojen perusteella (esim. Riihelä ym. 2007).

ei toteudu (mt. 49). Mittarin avulla ei voida kuitenkaan päätellä, kuinka köyhiä köyhyysrajan alapuolella olevat tulonsaajat todellisuudessa ovat.

Toinen tapa mitata köyhyyttä on köyhyyskuilu (engl. poverty gap ratio), joka huomioi samanaikaisesti sekä köyhyysrajan alapuolella olevien köyhien lukumäärän että sen, kuinka köyhiä he ovat. Köyhyyskuilu  $HI$  on päälukumitan  $H$  ja keskimääräisen köyhyyskuilun (engl. income gap ratio)  $I$  tulo<sup>12</sup>:

$$HI = \frac{1}{n} \sum_{i=1}^n [1 - y_i / z] \mathbb{I}(y_i \leq z) \quad (13)$$

Köyhyyskuilu on sitä suurempi, mitä suurempia köyhien väestöosuus ja keskimääräinen köyhyyskuilu ovat. Köyhyyskuilu (engl. poverty deficit) saadaan, kun köyhyysrajan alapuolelle olevien henkilöiden tulot vähennetään köyhyysrajasta ja lasketaan saadut erotukset yhteen. Keskimääräinen köyhyyskuilu [ $I = (1 - y_i / z)$ ] kuvaa tulojen suhteellista etäisyyttä köyhyysrajasta.

#### 2.1.4 Eriarvoisuus ja köyhyys Suomessa 1990-luvulla ja 2000-luvun alussa

Edellä selvitettiin tuloerojen ja köyhyyden laskemista erilaisten menetelmien avulla. Seuraavassa tarkastellaan, millaisia tuloerot ja köyhyys olivat Suomessa 1990-luvulla ja 2000-luvun alussa. Kappaleessa käsitellään myös alueiden sisäisten ja niiden välisten tuloerojen kehitystä 1990-luvun alun taloudellisen laman jälkeisistä vuosista aina vuoteen 2004 saakka. Alueluokituksena ovat maakunnat (ks. liite 1).

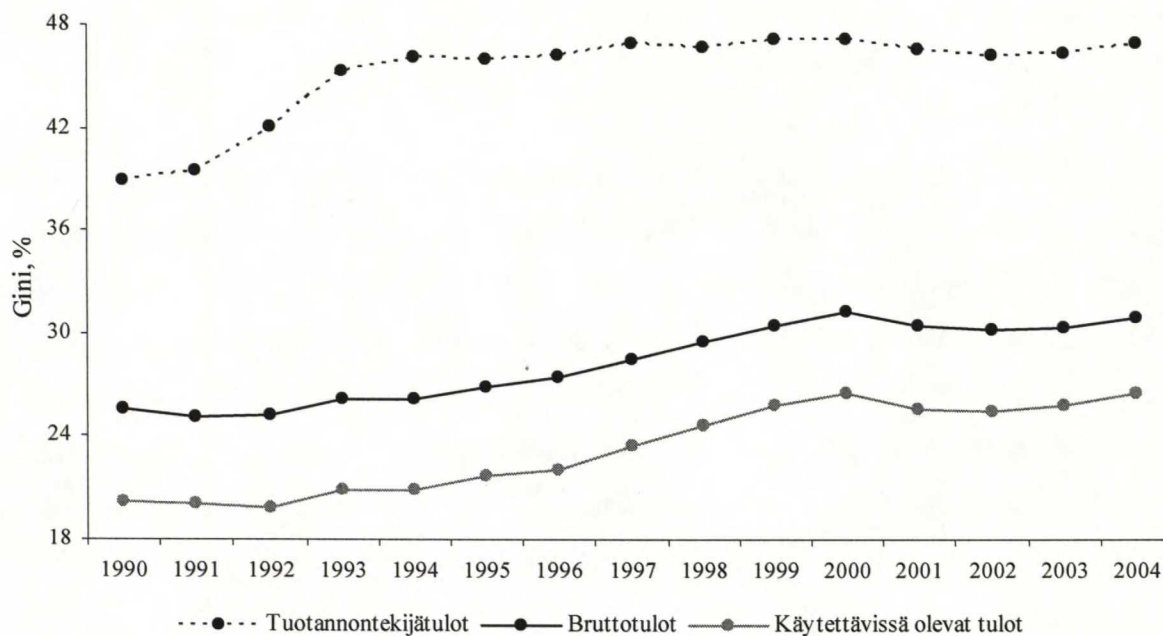
Kuvio 2 havainnollistaa tuloerojen kehitystä Suomessa vuosina 1990–2004. Siitä ilmenee, että Gini-kertoimella mitattu eriarvoisuus kasvoi voimakkaasti syvimmän laman aikaan vuosina 1992–1993. Bruttotulojen ja käytettävissä olevien tulojen Gini-kerroin säilyi kuitenkin vakaana vuosien 1990–1994 aikana. 1990-luvun alun talouden syvimmän laman aikana tuloerot säilyivät suhteellisen muuttumattomina Suomessa (Loikkanen ym. 2005, 118).

---

<sup>12</sup> Päälukumitta ja köyhyyskuilu ovat FGT-indeksin erityistapauksia (Foster, Greer & Thorbecke 1984).



Kuvio 2. Gini-kertoimella mitattu eriarvoisuus Suomessa vuosina 1990–2004 kolmella tulokäsitteellä (modifioitu OECD-skaala)



Lähde: Tilastokeskus 2006c, 13

Vuoden 1994 jälkeen talous alkoi kasvaa voimakkaasti. Kuviosta 2 ilmenee, että tuotannontekijätulojen Gini-kerroin pysyi vakaana. Bruttotuloissa ja käytettävissä olevissa tuloissa mitattu eriarvoisuus kasvoi kuitenkin voimakkaasti. Eriarvoisuuden kasvu jatkui aina vuosituhannen vaihteeseen asti, jolloin sen kasvu taittui. Eriarvoisuus pieneni vuonna 2001 ja pysyi lähes ennallaan vuonna 2002.

Gini-kertoimen arvot kääntyivät uudelleen nousuun vuosien 2003 ja 2004 aikana. Käytettävissä olevien tulojen Gini-kerroin oli 25,8 prosenttia vuonna 2003, kun se vuotta myöhemmin oli 0,7 prosenttiyksikköä korkeampi. Vuonna 2004 käytettävissä olevien tulojen Gini-kerroin kohosi samalle tasolle kuin vuonna 2000, jolloin tuloerot saavuttivat laman jälkeisen huipun. Myös tuotannontekijätuloilla ja bruttotuloilla mitattuna tuloerot kasvoivat vuonna 2004, mutta Gini-kertoimet jäivät arvoiltaan hieman alhaisemmiksi kuin vuonna 2000. Yleissääntönä on, että tuloerot ovat pienet Gini-kertoimen ollessa alle 25 prosenttia ja suuret, kun Gini-kerroin ylittää 30 prosenttia (Talousneuvosto 2001, 3).

Myös eriarvoisuus alueiden sisällä ja niiden välillä kasvoi laman jälkeisinä vuosina Suomessa. Vuosina 1995–2004 alueiden sisällä Gini-kertoimen arvot kasvoivat jokaisessa Suomen



maakunnassa<sup>13</sup>. Vuonna 1995 eriarvoisuus oli suurimmillaan Ahvenanmaan, Uudenmaan ja Itä-Uudenmaan maakunnassa ja pienimmillään Keski-Pohjanmaan, Kainuun ja Lapin maakunnassa. Koko maan eriarvoisuuden kasvun ohella myös eriarvoisuus maakuntien sisällä kasvoi selvästi 1990-luvun lopulla.

Vuosituhanen vaihteessa maakuntien sisällä Gini-kertoimen kasvu tasaantui. 2000-luvun alkuvuosina tuloerot pienenevät erityisesti niissä maakunnissa, joissa ne kasvoivat nopeimmin 1990-luvun jälkipuoliskolla (Ruotsalainen 2005, 24). Vuosina 2003 ja 2004 kotitalouksien väliset tuloerot kääntyivät uudelleen nousuun. Tällöin tuloerot supistuivat ainoastaan Uudenmaan ja Pohjois-Pohjanmaan maakunnissa. Vuosina 2003 ja 2004 tuloerot olivat suurimmat Uudenmaan, Ahvenanmaan sekä Itä-Uudenmaan maakunnassa ja pienimmät Keski-Pohjanmaan, Kainuun ja Lapin maakunnassa. Tuloerot olivat näin ollen suurimmillaan ja pienimmillään samoissa maakunnissa kuin lähes vuosikymmen aiemmin 1990-luvun puolivälissä.

Myös maakuntien välillä tuloerot kasvoivat laman jälkeisinä vuosina 1995–2004<sup>14</sup>. Alueiden välisiä tuleroja tarkastellaan vertaamalla alueiden keskiarvotuloja. Kotitalouksien kulutusyksikköä kohti laskettu käytettävissä oleva rahatulo kasvoi eniten Ahvenanmaan kotitalouksilla, joiden keskiarvolla mitattu kotitalouden kulutusyksikköä kohti laskettu käytettävissä oleva rahatulo oli vuonna 2004 reaalisesti melkein 45 prosenttia korkeampi kuin vuonna 1995. Toiseksi nopeinta tulojen kasvu oli Uudellamaalla ja Itä-Uudellamaalla. Heikointa tulokehitys oli puolestaan Itä- ja Pohjois-Suomen maakunnissa. Vuonna 2004 Kainuun ja Lapin maakunnissa asuvien kotitalouksien keskimääräiset reaalitytulot olivat vain reilut 20 prosenttia vuoden 1995 tulotasoa korkeammat.

Eriarvoisuuden ohella myös köyhyydessä tapahtui huomattavia muutoksia Suomessa vuosina 1990–2004 (Riihelä ym. 2007, 13). Kuvio 3 havainnollistaa köyhyysasteen kehitystä Suomessa vuosina 1990–2004. Köyhyyttä mitataan kuviossa päälukumitalla ja kotitalouden käytettävissä olevat tulot on muunnettu henkilömääräisiksi modifioitun OECD-skaalan avulla. Köyhyysraja on 60 % mediaanitulosta ja tulojen keskiarvosta. Kuvioista 3 ilmenee, että köyhyysaste laski Suomessa 1990-luvun alun taloudellisen laman aikana, mutta nousi

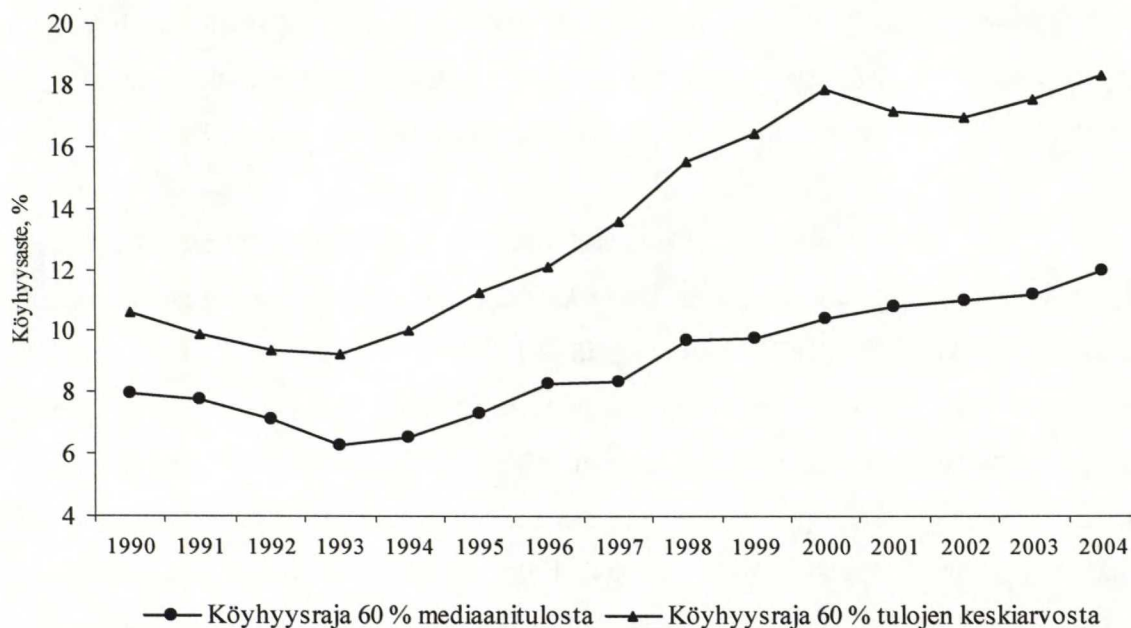
---

<sup>13</sup> Tiedot perustuvat Tilastokeskuksen (Tilastokeskus 2006b) laskelmiin. Eriarvoisuutta mitattiin käytettävissä olevista rahatuloista lasketulla Gini-kertoimella. Ekvivalenssiskaalana oli modifioitu OECD-skaala.

<sup>14</sup> Tiedot perustuvat Tilastokeskuksen (Tilastokeskus 2006a) laskelmiin.

voimakkaasti vuosikymmenen puolivälistä lähtien päätyen vuonna 2004 merkittävästi vuotta 1990 korkeammalle tasolle. Köyhyysasteen kasvu merkitsee, että köyhyysrajan alapuolella olevien henkilöiden kokonaismäärä kasvoi Suomessa vuodesta 1990 vuoteen 2004.

Kuvio 3. Päälukumitalla mitatun köyhyysasteen kehitys Suomessa vuosina 1990–2004 (modifioitu OECD-skaala)



Lähde: Riihelä ym. 2007, 27

Vuosina 1995–2004 köyhien lukumäärän kasvun lisäksi myös köyhyys syveni (Riihelä ym. 2007, 22). Köyhyyskuilun kasvu osoittaa, että köyhyysrajan alapuolella olevien henkilöiden kokonaisköyhyyskuilu kasvoi vuosina 1995–2004 (ks. Riihelä ym. 2007, 27). Riihelä ym. (2007, 15) havaitsivat, että köyhyys lisääntyi Suomessa vuosina 1990–2004 riippumatta valitusta köyhyysmittarista tai köyhyysrajasta.

## 2.2 Väestön terveys

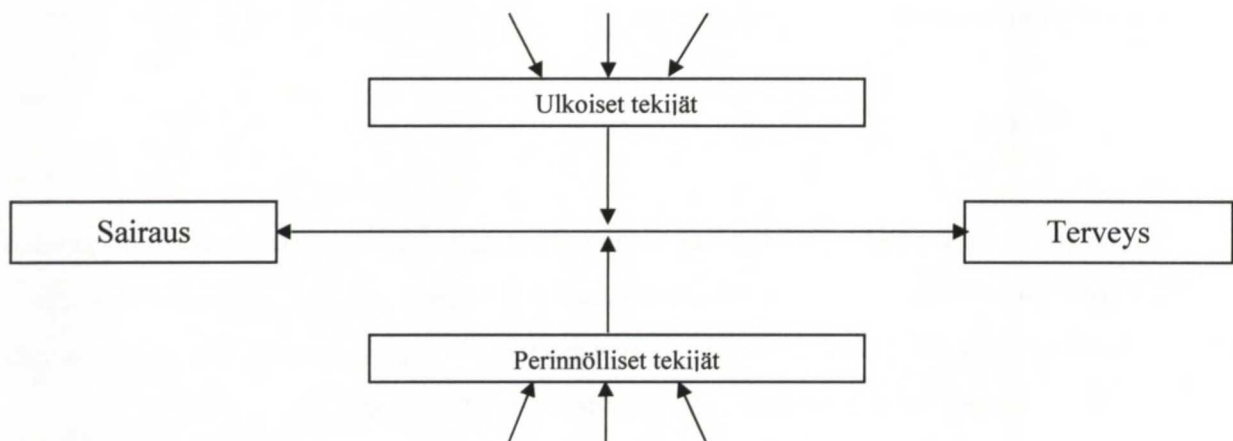
Seuraavassa tarkastellaan väestön terveyttä Suomessa. Aluksi selvitetään terveyden määritelmää ja erilaisia terveyden mittausmenetelmiä. Tämän jälkeen tarkastellaan väestön terveydentilaa Suomessa kuvattujen terveystietojen näkökulmasta. Luvun lopussa selvitetään terveyden sosioekonomisia ja alueellisia eroja Suomessa.

### 2.2.1 Terveiden määritelmästä ja mittaamisesta

Terveys on moniulotteinen käsite ja sen arvioiminen on ollut monien tutkijoiden pohdinnan aiheena (Skalli ym. 2006, 33). Maailman terveysjärjestö (WHO) on lausunut aikanaan kenties maailman tunnetuimman terveyden määritelmän. WHO (1948) määrittelee terveyden täydellisenä fyysisen, psyykkisen ja sosiaalisen hyvinvoinnin tilana. Terveyttä voidaan tarkastella myös yhtenä elintason mittarina, sillä terveys on tärkeä osa yksilöiden hyvinvointia (Lyngstadt ym. 1997, 10). Terveydellä on välineellinen arvo, sillä se on edellytys esimerkiksi menestyksekkäälle työnteolle, mutta terveys on jo arvo sinänsä (Deaton 1999a, 1).

Yksilön terveys ja hyvinvointi ovat yhteydessä lukuisiin eri tekijöihin. Kuviossa 4 terveyteen ja hyvinvointiin vaikuttavat tekijät on jaettu kahteen luokkaan: perinnöllisiin ja luontaisiin sekä ulkoisiin tekijöihin. Yksilön perinnöllisiä ja luontaisia tekijöitä ovat muun muassa ikä, perimä ja sukupuoli. Ulkoisia tekijöitä ovat puolestaan esimerkiksi sosiaaliluokka, ammatti, koulutus, elintavat, työ- ja elinolot, sosiaaliset suhteet sekä ympäristö. (Euroopan komissio 2007, 2) Yksilön koulutus ja ammatti yhdessä tulojen, omaisuuden ja asumistason kanssa mittaavat yksilön sosioekonomista asemaa, jolla on todettu olevan yhteys terveyteen (Martelin ym. 2005, 271). Sosiaaliset olosuhteet vaikuttavat merkittävästi yksilön mahdollisuuteen elää terveenä. Köyhyys, syrjäytyminen, huonot asuinolot, vaikeat lapsuuden elinolosuhteet ja alhainen ammattiasema ovat monia sairauksia, kuolemia sekä maiden sisäisiä ja niiden välisiä terveyseroja määrittäviä tekijöitä. (World Health Organization 2004, 1)

Kuvio 4. Terveysteen ja hyvinvointiin vaikuttavia tekijöitä



Lähde: Euroopan Komissio 2007, 2



Terveyden moniulotteisuudesta johtuen lukuisia erilaisia mittareita on kehitetty kuvaamaan terveyttä sekä väestö- että yksilötasolla (Euroopan komissio 2007, 2). Karkeasti nämä mittarit voidaan jakaa makro- ja mikrotason indikaattoreihin. Makrotason terveystmittarit antavat kokonais kuvan terveydestä väestötasolla. Elinajanodote, kuolleisuusluvut ja lapsikuolleisuus ovat esimerkkejä makrotason terveystindikaattoreista. Mikrotason terveystmittarit kuvaavat puolestaan terveydentilaa yksilötasolla. Mikrotason indikaattoreita ovat sekä objektiiviset mittarit, kuten verenpaine, näkö tai kuulo että subjektiiviseen arviointiin perustuvat terveydentilan mittarit. Yhtenä yleisimpänä subjektiiviseen arviointiin perustuvana mikrotason terveystmittarina käytetään kyselyä, jossa haastateltavia pyydetään määrittelemään heidän nykyinen terveydentilansa jonkin hierarkkisen asteikon perusteella (esim. Miten kuvaisitte nykyistä terveydentilaanne? Sanoisitteko, että se on erinomainen, todella hyvä, hyvä, kohtalainen vai huono?). (mt. 2-3) Vaikka mittari perustuu subjektiiviseen arviointiin, on sen todettu ennustavan tutkimuksissa hyvin muuta terveydentilaa ja terveyttä tulevaisuudessa (Zimmer ym. 2000, 456).

#### 2.2.2 Terveys ja sen väestöryhmittäiset erot Suomessa

Makrotason terveystmittarit, kuten elinajanodote ja kuolleisuus antavat kokonais kuvan väestön terveydestä. Suomessa elinajanodote on noussut nopeasti viime vuosikymmeninä. Toisen maailmansodan ja 1950-luvun alun välisen kasvuvaiheen jälkeen elinajanodote on kasvanut uudelleen nopeasti vuodesta 1970 lähtien (Martelin ym. 2005, 117). Ennen 1950-lukua elinajanodotteen kasvuun vaikuttivat merkittävästi tartuntatauti- ja lapsikuolleisuuden pieneneminen. Viime vuosikymmeninä kroonisten tautien, kuten sepelvaltimo- ja muiden verenkiertoelinsairauksien aiheuttaman kuolleisuuden pieneneminen on vaikuttanut merkittävästi elinajanodotteen kasvuun. (ma. 117) Suomessa miesten elinajanodote oli 75,3 vuotta ja naisten 82,3 vuotta vuonna 2004 (Martelin ym. 2006, 48).

1970-luvun alkupuolelta lähtien kuolleisuuden pieneneminen on ollut suhteellisesti nopeinta imeväisiässä. Suomessa imeväiskuolleisuus on maailman pienimpiä ja lapsuusiän kuolleisuus on vähentynyt kansainvälisesti hyvin alhaiselle tasolle. (Martelin ym. 2005, 117) Myös vanhusten kuolleisuus on vähentynyt Suomessa merkittävästi viimeisen kolmen vuosikymmenen aikana. Keski-ikäisen väestön kuolleisuus on pienentynyt melkein puolella vastaavana ajanjaksona pääasiassa verenkiertoelinten sairauksien vähenemisen takia.

(Martelin ym. 2006, 48) Myös muissa OECD-maissa keski-ikäisen väestön, erityisesti 50–64-vuotiaiden miesten kohdalla pienentyneen kuolleisuuden taustalla on voimakas verenkiertoelinsairauksien aiheuttaman kuolleisuuden pieneneminen (Deaton 2004, 26).

Suomalaisten omaa arviota terveydentilasta ja pitkäaikaisista sairauksista on selvitetty kysely- ja haastattelututkimuksin 1960-luvulta alkaen. Kaksi kolmasosaa suomalaisesta aikuisväestöstä arvioi terveydentilansa hyväksi tai melko hyväksi. (Manderbacka 2005, 130–131). Keskinkertaiseksi tai huonoksi terveytensä arvioi noin yksi kolmasosa suomalaisista 2000-luvun alussa. Suomalaisen aikuisväestön terveydentila on parantunut jossakin määrin vuosien 1979 ja 2002 välisenä aikana. Koettu terveys parani erityisesti lamavuosien aikana vuosina 1992–1994. (Rahkonen ym. 2004, 2161) Pitkäaikaissairauksien osalta kehityssuunta on ollut päinvastainen verrattuna koettuun terveyteen. Väestön koetun terveyden parantuessa erityisesti lievien pitkäaikaissairauksien määrä on lisääntynyt (Manderbacka 2005, 133). Arviolta puolet suomalaisista aikuisista ilmoittaa sairastavansa ainakin yhtä pitkäaikaista sairautta (ma. 131).

Viime vuosien suotuisasta kehityksestä huolimatta Suomessa väestöryhmien välillä on suuria eroja useilla eri terveystilanteilla mitattuina (Martelin ym. 2005, 266). Väestöryhmien terveyseroja voidaan tarkastella useiden eri tekijöiden, kuten esimerkiksi sukupuolen, siviilisäädyn, koulutuksen, tulojen ja alueen perusteella. Seuraavassa selvitetään tarkemmin väestöryhmien sosioekonomisia ja alueellisia terveyseroja Suomessa. Alueluokituksena ovat maakunnat.

Terveyserot ihmiskunnan hyvinvoivan eliitin ja huonoimmassa asemassa olevien köyhien välillä ovat tällä hetkellä historian suurimmat (Vuorinen 2007, 89). Lukuisilla sosioekonomisen aseman osoittimilla, kuten esimerkiksi tuloilla, koulutuksella ja ammatilla on todettu olevan selvä yhteys terveyteen. Vaikka elinajanodote on kasvanut kaikissa sosioekonomisissa ryhmissä, ryhmien väliset erot kuolleisuudessa ovat Suomessa huomattavat moniin muihin maihin verrattuna. Viime vuosikymmeninä kuolleisuuserot ovat jopa kasvaneet. Esimerkiksi 1990-luvun lopulla ylempiin toimihenkilöihin kuuluvalla 35-vuotiaalla miehellä oli odotettavissa elinaikaa kuusi vuotta enemmän kuin samanikäisellä työntekijämiehellä. Naisilla ero oli 3,2 vuotta. Myös koulutuksen ja tulojen mukaan tarkasteltuna elinajanodotteen erot ovat samankaltaisia. (Martelin ym. 2005, 271–273)



Kuolleisuuden ohella myös sairastavuuserot sosioekonomisten ryhmien välillä ovat merkittävät Suomessa. Esimerkiksi verenkierto-, tuki- ja liikuntaelinten sairauksissa sekä huonoksi koetussa terveydessä sosioekonomisten ryhmien väliset erot vastaavat suuruudeltaan näiden ryhmien kuolleisuuseroja. Suurten sairastavuuserojen takia sosioekonomisten ryhmien välillä erot terveisinä eletyissä elinvuosissa ovat elinajanodote-eroja vieläkin jyrkemmät. (Martelin ym. 2005, 272)

Sosioekonomisten ryhmien kuolleisuuserojen ohella myös elinajanodotteen alueelliset erot ovat Suomessa merkittävät. Maakunnittain tarkasteltuna elinajanodote oli korkea Pohjanmaalla, jossa miesten odotettavissa oleva elinaika oli 76,8 vuotta ja naisilla 82,6 vuotta vuosina 1997–2001. Myös Ahvenanmaalla elinajanodote oli samaa luokkaa kuin Pohjanmaalla. Suomessa elinajanodotteen toista ääripäätä edustavat maakunnista Kymenlaakso, Pohjois-Karjala, Kainuu ja Lappi, missä elinajanodote oli miehillä neljä vuotta ja naisilla 2-3 vuotta Pohjanmaata ja Ahvenanmaata lyhyempi. (Martelin ym. 2005, 270)

Myös muilla mittareilla terveyden alue-erot ovat samansuuntaisia. Esimerkiksi pitkäaikaissairaudet olivat Itä-Suomessa muuta maata yleisempiä. (Martelin ym. 2005, 270) 2000-luvun alussa koetun terveyden osalta alueelliset erot olivat varsin pienet, mutta merkkejä itäsuomalaisten maan muita osia huonommasta terveydestä oli edelleen nähtävissä (Manderbacka 2005, 131). Myös Helakorpi ym. (2007, 12) havaitsivat, että hyväksi koetussa terveydentilassa ei ollut suuria eroja maakuntien välillä. Pääkaupunkiseudulla ja Pohjanmaalla terveydentilansa hyväksi tai melko hyväksi kokevien osuus oli kuitenkin tutkimuksen mukaan hieman maan keskitasoa suurempi vuosina 1978–2005. Myös itäuusimaalaisten ja kymenlaaksolaisten naisten keskuudessa terveytensä hyväksi kokevien osuus oli hieman suurempi kuin muun maan naisväestössä. Terveytensä hyväksi kokevien osuudet olivat puolestaan hieman maan keskitasoa alhaisempia muun muassa pohjoiskarjalaisten naisten ja eteläsavolaisten miesten keskuudessa. (mt. 12)

Kansaneläkelaitoksen (2007) Terveyspuntarissa julkaistaan vuosittain tietoja Suomen väestön terveydestä alueittain. Kelan rekisteritietojen avulla eri aluejaotuksilla laskettu sairastavuusindeksi kuvaa, kuinka sairasta tai tervettä kyseessä olevan alueen väestö on suhteessa koko maan keskiarvoon. Ikävakioidun sairastavuusindeksin perusteella ihmiset olivat kaikkein terveimpiä Ahvenanmaan maakunnassa tarkasteltaessa vuosia 2003 ja 2004. Ahvenanmaan jälkeen väestö oli terveintä sairastavuusindeksin perusteella Uudenmaan, Itä-



Uudenmaan ja Pohjanmaan maakunnissa tällöin. Kaikkein sairainta väestö oli puolestaan vastaavalla ajanjaksolla Kainuun maakunnassa, jonka jälkeen tulivat Pohjois-Savo ja Pohjois-Karjala. Edellä esitettyjen elinajanodotteen ja pitkäaikaissairastavuuden lisäksi myös sairastavuusindeksin perusteella voidaan havaita, että maan itäosissa väestön terveys on Etelä- ja Länsi-Suomea huonompi.

### 3 TULOEROJEN JA TERVEYDEN VÄLISEN YHTEYDEN TEOREETTINEN TARKASTELU

#### 3.1 Yleistä

Kuten edellä havaittiin, terveyteen vaikuttavat useat eri tekijät. Yksilön luontaisten tekijöiden, kuten perimän tai sukupuolen ohella myös ulkoiset tekijät vaikuttavat terveyteen. Toisaalta terveys voi olla yhteydessä esimerkiksi sellaisiin ulkoisiin tekijöihin, kuten tuloihin, sillä hyvä terveys on edellytys menestyksekkäälle työnteolle. Tarkasteltaessa tulojen ja terveyden välisen yhteyden mekanismeja on huomioitava, että tulojen ja terveyden yhteys voi toimia kahteen suuntaan: joko tulot parantavat terveyttä tai hyvä terveys lisää yksilön tuloja esimerkiksi hyvän työkyvyn myötä (Deaton 2002, 15). Toisaalta sekä tulot että terveys voivat olla korreloituneet muiden tekijöiden kanssa (ma. 15). Seuraavassa esitetään katsaus<sup>15</sup> hypoteeseista, joita tulojen, köyhyyden ja tuloerojen yhteydestä terveyteen on kirjallisuudessa esitetty.

Absoluuttisen tulon hypoteesin (engl. absolute income hypothesis) mukaan tulot ovat yhteydessä terveyteen, mutta tuloerot eivät vaikuta siihen (Deaton 2003, 114). Esimerkiksi Preston (1975) havaitsi aikoinaan, että tulotason nousu nosti elinajanodotetta enemmän köyhissä kuin rikkaissa maissa. Tulojen ja kuolleisuuden välinen yhteys oli epälineaarinen siten, että tulojen elinajanodotetta nostava vaikutus pieneni tulotason noustessa. On esitetty, että tulot vaikuttavat terveyteen, koska tulotason nousu lisää yksilöiden terveyteen vaikuttavien hyödykkeiden kulutusmahdollisuuksia (Gravelle 1999, 101). Absoluuttisen tulon hypoteesin mukaan terveys on sitä parempi, mitä suuremmat tulot yksilöllä on. Yhteys on hypoteesin mukaan konkaavi eli tulojen terveyttä lisäävä vaikutus pienenee tulotason noustessa. (Li & Zhu 2006, 3)

---

<sup>15</sup> Ks. myös Wagstaff & van Doorslaer (2000, 545–550).

Köyhyyshypoteesi (engl. poverty hypothesis) on verrattavissa absoluuttisen tulon hypoteesiin, sillä huono terveys voi olla seurasta alhaisesta tulotasosta tai äärimmäisestä köyhyydestä (Li & Zhu 2006, 3). Toisaalta köyhyyshypoteesin mukaan yksilön tulot eivät sellaisenaan määritä yksilön terveydentilaa, vaan terveys on yhteydessä köyhyyskuiluun eli köyhyysrajan ja yksilön tulojen erotuksena mitattuun köyhyyteen (Wagstaff & van Doorslaer 2000, 548). Hypoteesin mukaan terveys on sitä huonompi, mitä köyhempi yksilö on.

Suhteellisen tulon hypoteesin (engl. relative income hypothesis) mukaan terveys riippuu yksilön tuloista suhteessa muiden tuloihin samassa ryhmässä. Hypoteesin mukaan terveys heikkenee, kun yksilö on taloudellisesti huono-osainen suhteessa viiteryhmäänsä ja päinvastoin. (Li & Zhu 2006, 5) Suhteellisen tulon laskemisessa yksilön tulojen vertailu kohdistetaan siihen viiteryhmään, johon yksilö kuuluu (Miller & Paxson 2001, 5). Wilkinsonin (1997, 591) mukaan kehittyneissä maissa suhteelliset tulot vaikuttavat terveyteen enemmän kuin absoluuttinen tulotaso. Suhteellisen tulon hypoteesi on verrattavissa suhteellisen aseman hypoteesiin (engl. relative position hypothesis), jonka mukaan terveyteen vaikuttavat tulojen lisäksi yksilön asema koko tulojakaumassa. Suhteellista asemaa voidaan mitata tulojen lisäksi myös esimerkiksi ammatilla tai koulutuksella. (Li & Zhu 2006, 5)

Edellä esitetyissä hypoteeseissa tuloerojen ja terveyden välillä ei oletettu vallitsevan yhteyttä. Seuraavassa käsitellään tulojen eriarvoisuuden hypoteesia (engl. income inequality hypothesis). Kyseessä olevan hypoteesin mukaan tuloerot vaikuttavat haitallisesti yksilöiden terveyteen (Wagstaff & van Doorslaer 2000, 549). Hypoteesi väittää, että tuloerot ovat sellaisenaan uhka yksilöiden terveydelle (Li & Zhu 2006, 3). Seuraavassa esitetään malli tuloerojen ja terveyden välisestä yhteydestä ja tarkastellaan yksityiskohtaisemmin tulojen eriarvoisuuden hypoteesia.

### 3.2 Tulojen eriarvoisuuden hypoteesi

Deaton (1999a) on esittänyt mallin tuloerojen ja terveyden välisestä yhteydestä. Tarkastelu perustuu suhteellisen tulon malliin, joka lähtee oletuksesta, että yksilön terveys on riippuvainen yksilön asemasta viiteryhmässä (mt. 12). Deaton (1999a, 16–17) laajentaa suhteellisen tulon mallia siten, että viiteryhmän sisäisillä tuloeroilla on suora vaikutus yksilön terveyteen. Mallissa yksilön terveys määritellään viiteryhmän suurituloisimman henkilön ja tarkastelun kohteena olevan yksilön tulojen erotuksen suhteen vähenevänä funktiona:



$$H = \alpha - \beta(y^{\mu}_{\max} - y), \quad (14)$$

jossa  $H$  on yksilön terveys,  $y$  yksilön tulot ja  $y^{\mu}_{\max}$  viiteryhmän suurituloisimman henkilön tulot. Parametri  $\alpha$  on viiteryhmän suurituloisimman henkilön terveys ja parametri  $\beta$  kuvaa sitä vaikutusta, joka aiheutuu yksilön terveydelle absoluuttisesta tuloerosta viiteryhmän rikkaimman ja kyseessä olevan henkilön välillä.

Viiteryhmän suurituloisimman henkilön tulot ovat  $\theta$  keskihajontaa viiteryhmän keskiarvotulon yläpuolella siten, että:

$$y^{\mu}_{\max} = \mu + \theta\sigma_{\varepsilon}, \quad (15)$$

jossa  $\mu$  on viiteryhmän keskimääräinen tulotaso ja  $\sigma_{\varepsilon}$  viiteryhmän sisäiset tuloerot.

Sijoitettaessa yhtälö 15 yhtälöön 14 voidaan yksilön terveys esittää funktiona yksilön tulojen, viiteryhmän keskiarvotulon ja tuloerojen suhteen:

$$H = \alpha + \beta y - \beta \mu - \beta \theta \sigma_{\varepsilon}. \quad (16)$$

Yhtälössä 16 viiteryhmän sisäisten tuloerojen kasvaessa yksilön terveys huononee. Tulojen eriarvoisuuden hypoteesin mukaisesti tuloerot ovat uhka sellaisenaan yksilöiden terveydelle. Yhtälössä 16 myös viiteryhmän keskimääräisen tulon kasvu heikentää yksilön terveyttä. Suhteellisen tulon hypoteesin mukaan yksilön terveys heikkenee, kun viiteryhmän keskimääräinen tulotaso kasvaa yksilön tulojen säilyessä ennallaan (Gerdtham & Johannesson 2004, 234). Yksilön tulotason nousu puolestaan parantaa terveyttä. Tämä on yhtenevä edellä esitetyn absoluuttisen tulon hypoteesin kanssa. Suurituloisimman terveys ei kuitenkaan parane kyseessä olevan henkilön tulojen noustessa, sillä terveyteen vaikuttaa yhtälön 14 mukaan tulokuilu rikkaimman henkilön ja muiden viiteryhmän henkilöiden välillä (Deaton 1999a, 17). Tulojakaumassa suurituloisimman henkilön alapuolella olevien yksilöiden terveys kuitenkin heikkenee, kun rikkaimman henkilön tulot nousevat ja tuloerot kasvavat. On huomattava, että suurituloisimman henkilön valinta vertailukohteeksi on täysin sattumanvaraista. Myös



pienituloisin yksilö voisi olla vertailukohteena. Tällöin tuloerojen kasvulla olisi terveyttä suojaava vaikutus ja yhtälön 16 viimeisen termin etumerkki olisi positiivinen. (ma. 17)

Edellä esitetty malli jättää huomioimatta kolme tekijää tulojen ja terveyden välisestä yhteydestä. Ensinnäkin mallin oletus on, että tulot vaikuttavat terveyteen. Se ei siten huomioi mahdollista käänteistä kausaalisuutta tulojen ja terveyden välillä. Toiseksi mallissa tarkastellaan suhteellisten tulojen ja terveyden välistä yhteyttä, eikä siinä huomioida muita yksilön terveyteen vaikuttavia tekijöitä. (Deaton 1999a, 30) Näitä tekijöitä tarkasteltiin aiemmin luvussa 2.2.1. Kolmantena huomioitavana tekijänä on mallin oletus tulojen ja terveyden välisen yhteyden lineaarisuudesta (Deaton 1999a, 30). Oletus poikkeaa tulojen ja terveyden välisen yhteyden konkaavisuusolettamasta, jota tarkastellaan jäljempänä.

Tuloerojen ja terveyden välistä yhteyttä tarkastelleissa tutkimuksissa on pyritty selvittämään, vaikuttavatko tuloerot eri väestöryhmien terveyteen samalla tavalla (Dahl ym. 2006, 2563). Tulojen eriarvoisuuden hypoteesista on esitetty kaksi hieman toisistaan poikkeavaa versiota. Tulojen eriarvoisuuden hypoteesin niin sanotun vahvan version mukaan tuloerot ovat haitallisia kaikkien yksilöiden terveydelle yksilön tulotasosta riippumatta. Tulojen eriarvoisuuden niin sanotun heikon version mukaan tuloerot ovat haitallisimpia yhteiskunnan huono-osaisten terveydelle. (Mellor & Milyo 2002, 513–514) Hypoteesin heikko versio on rinnastettavissa suhteellisen tulon hypoteesiin, vaikka hypoteesit eivät täysin vastaakaan toisiaan. Suhteellisesti tarkasteltuna eriarvoisuuden kasvu merkitsee, että tulojen ero rikkaimman ja köyhimmän välillä kasvaa. Tällöin tuloerojen haitallinen vaikutus kohdistuu voimakkaammin tulojen suhteen huono-osaisiin. Tulojen eriarvoisuuden hypoteesin vahvan version mukaan myös yhteiskunnan varakkaimmat yksilöt voivat kärsiä tuloerojen haitallisista vaikutuksista terveydelle. (Li & Zhu 2006, 6) Haitallinen vaikutus voi olla seurausta esimerkiksi tuloerojen kasvun myötä lisääntyvästä rikollisuudesta.

Tässä tutkielmassa testataan empiirisesti tulojen eriarvoisuuden hypoteesin vahvaa ja heikkoa versiota. Tulojen eriarvoisuuden hypoteesin vahva versio määritellään:

$$H_{ij} = \beta_0 + \beta_1 Q_j + I_{ij} \Gamma + X_{ij} \Theta + \varepsilon_{ij} , \quad (17)$$

joka on muunnelma Lin ja Zhun (2006, 5) esityksestä. Yhtälössä 17 indeksi  $i$  viittaa yksilöön ja  $j$  yhteisöön (engl. community).  $H_{ij}$  kuvaa yksilön  $i$  terveydentilaa yhteisössä  $j$ .  $Q_j$  tarkoittaa tuloeroja yhteisössä  $j$ .  $I_{ij}$  on yksilön tulojen sekä tulojen neliön vektori ja parametrivektori  $\Gamma = \{\Gamma_1, \Gamma_2\}$ .  $X_{ij}$  on puolestaan vektori, joka kuvaa muita yksilön ja yhteisön ominaisuuksia. Parametrivektori  $\Theta = \{\Theta_1, \Theta_2, \dots, \Theta_Z\}$ , jossa  $Z$  on yksilön ja yhteisön ominaisuuksia kuvaavien taustamuuttujien lukumäärä. Käsillä olevassa tutkielmassa yksilön ja yhteisön ominaisuuksia kuvaaviin taustamuuttujiin sisällytetään muun muassa yhteisön  $j$  keskimääräinen tulotaso edellä esitetyn Deatonin (1999a) malliin perustuen. Yhteisö  $j$  on yksilön  $i$  viiteryhmä, johon vertailu kohdistetaan. Tulojen eriarvoisuuden hypoteesin vahvan version mukaisesti oletetaan, että parametri  $\beta_1 < 0$ . Näin ollen terveydentila heikkenee tuloerojen kasvaessa.

Tarkasteltaessa tulojen eriarvoisuuden hypoteesin heikkoa versiota laajennetaan edellä esitettyä yhtälöä 17 siten, että tuloerojen vaikutus terveyteen vaihtelee yksilön suhteellisen tulotason mukaan. Tätä varten yhtälöön 17 lisätään muuttuja  $Q_j R_{ij}$ , jossa  $Q_j$  tarkoittaa tuloeroja yhteisössä  $j$ . Yksilön  $i$  suhteellista tulotasoa merkitään puolestaan muuttujalla  $R_{ij}$ . Suhteellinen tulotaso määritellään sen perusteella, mikä on yksilön asema yhteisön  $j$  tulojakaumassa. Mitä korkeampi yksilön absoluuttinen tulotaso on, sitä suurempia arvoja  $R_{ij}$  sekä  $Q_j R_{ij}$  saavat. (Li & Zhu 2006, 5) Lin ja Zhun (2006, 5) esitystä mukaillen tulojen eriarvoisuuden hypoteesin heikko versio määritellään:

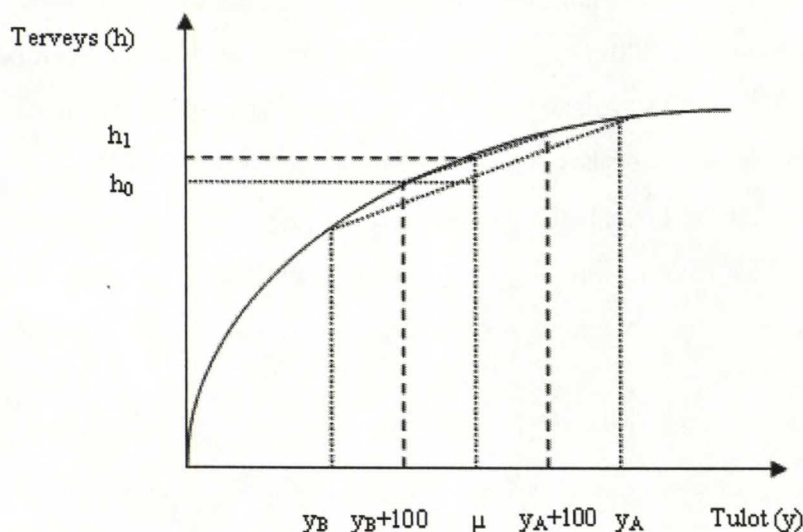
$$H_{ij} = \beta_0 + \beta_1 Q_j + \delta R_{ij} + \eta Q_j R_{ij} + I_{ij} \Gamma + X_{ij} \Theta + \varepsilon_{ij}. \quad (18)$$

Yhtälössä 18 muut muuttujat määritellään kuten yhtälössä 17. Tulojen eriarvoisuuden hypoteesin heikon version mukaisesti oletetaan, että parametri  $\eta$  saa positiivisia arvoja. Tuloerojen haitallinen vaikutus terveydelle on näin ollen sitä pienempi, mitä korkeampi yksilön suhteellinen tulotaso on eli mitä suurempia arvoja muuttuja  $R_{ij}$  saa. (Li & Zhu 2006, 5) Oletuksena on tässä, että terveys on sitä parempi, mitä suurempia arvoja  $H_{ij}$  saa.



Kolme vuosikymmentä sitten tutkijat havaitsivat, että väestötasolla tuloeroilla oli yhteys terveyteen, kun sitä mitattiin esimerkiksi elinajanodotteella tai lapsikuolleisuudella. Mikäli tulojen ja terveyden välisen yhteyden oletetaan olevan yksilötasolla konkaavi, väestön keskimääräinen terveys paranee keskimääräisen tulotason noustessa ja tuloerojen pienentyessä. Tilannetta havainnollistetaan kuviossa 5. Kuviossa esitetty funktio kuvaa terveyden  $h$  tuotantofunktiota tulojen  $y$  suhteen (ks. Wagstaff & van Doorslaer 2000, 545–546). Kuvitteellisen väestön muodostavat henkilöt A ja B, joiden tulot ovat  $y_A$  ja  $y_B$  siten, että  $y_A > y_B$ . Väestön keskimääräinen tulotaso on  $\mu$  ja terveys  $h_0$ . Siirrettäessä 100 rahayksikköä henkilöltä A henkilölle B tuloterot pienenevät ja väestön keskimääräinen terveys paranee arvosta  $h_0$  arvoon  $h_1$ . Absoluuttisen tulon hypoteesin mukaan tuloerojen ja terveyden välinen yhteys väestötasolla on seurausta tulojen ja terveyden konkaavista yhteydestä yksilötasolla. Tästä johtuen tuloerojen ja väestön keskimääräisen terveyden välisestä yhteydestä on käytetty nimitystä konkaavisuusvaikutus (engl. concavity effect) (Subramanian & Kawachi 2004, 80). On esitetty, että tuloerojen ja väestön terveyden välillä havaittu yhteys olisi näin ollen tulojen ja terveyden välisestä yhteydestä aiheutuva virhepäättelmä (Gravelle 1999, 100).

Kuvio 5. Tulojen ja terveyden välisen yhteyden konkaavisuuden seuraukset



Lähde: Wagstaff & van Doorslaer 2000, 546

Tuloerojen ja terveyden välisestä yhteydestä ei voida tehdä päätelmiä pelkästään väestötasolla, sillä tuloerojen ja väestön keskimääräisen terveydentilan välinen yhteys voi olla seurausta joko tulojen ja terveyden välisestä yhteydestä yksilötasolla tai tuloerojen



haitallisuudesta kaikkien yksilöiden terveydelle sellaisenaan (Subramanian & Kawachi 2004, 78). Jälkimmäisessä tapauksessa tuloeroilla on suora vaikutus väestön keskimääräiseen terveyteen siten, että tuloista riippumatta yksilöiden terveys on huonompi tulonjaoltaan eriarvoisessa yhteisössä kuin mitä se olisi tulojen jakautuessa tasaisesti. Konkaavisuusvaikutuksesta poiketen ilmiötä on kutsuttu kontekstuaaliseksi vaikutukseksi (engl. contextual effect). Tuloerojen kasvaessa kontekstuaalisuusvaikutus näkyisi terveyden tuotantofunktion  $h$  siirtymisenä alaspäin kuviossa 5. (ma. 80)

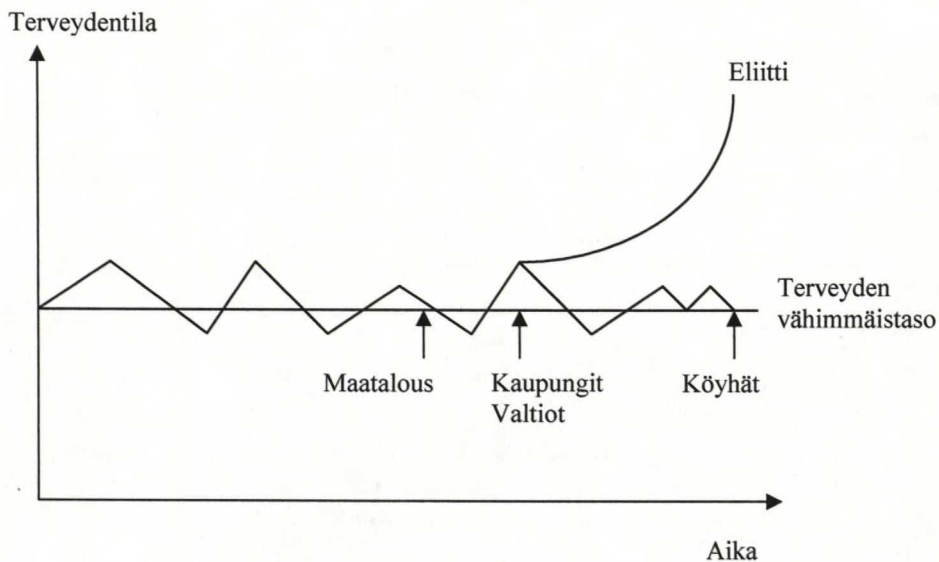
Tulojen eriarvoisuuden hypoteesin testaaminen vaatii sekä konkaavisuusvaikutuksen että kontekstuaalisen vaikutuksen yhtäaikaista mittaamista. Tämän vuoksi tulojen eriarvoisuuden hypoteesin testaamisessa tarvitaan monitasoinen aineisto (engl. multilevel data), jossa on tietoja sekä yksilöiden tuloista ja terveydestä että viiteryhmän tuloeroista. (Subramanian & Kawachi 2004, 80)

### 3.3 Vaikutusmekanismeja tuloerojen ja terveyden välillä

Tulojen eriarvoisuuden hypoteesin mukaan tuloerot ovat haitallisia yksilöiden terveydelle. Väitteet ihmisten välisen eriarvoisuuden haitallisuudesta juontavat juurensa metsästys-keräilykulttuurin aikakauteen (Deaton 2001, 4–5). Metsästys-keräilykulttuurissa yhteisöt olivat monissa asioissa suhteellisen tasa-arvoisia verrattuna esimerkiksi niitä seuranneeseen maanviljelykulttuurin. On esitetty, että metsästys-keräilykaudella ihmismielen kehittyminen mahdollisti ihmisten välisen yhteistyön ja sosiaalisten suhteiden muodostumisen, jotka olivat välttämättömiä yhteisöjen eloonjäämisessä. Ihmislaajalle kehittynyt mieltymys tasa-arvoon ja oikeudenmukaisuuteen oli tehokas keino yhteisöjen menestykseen ja hyvinvointiin metsästys-keräilykulttuurissa. (mt. 4–5)

Sosiaalisten hierarkioiden syntyessä ensimmäisten korkeakulttuurien aikaan noin 3000–500 ennen ajanlaskun alkua myös terveys alkoi jakautua sosiaalisten hierarkioiden mukaisesti (Vuorinen 2007, 90). Kuvio 6 havainnollistaa terveydentilan eriytymisen kehitystä hierarkkisten yhteiskuntien välillä ja sisällä.

Kuvio 6. Terveydentilan eriytyminen hierarkkisten yhteiskuntien välillä ja sisällä



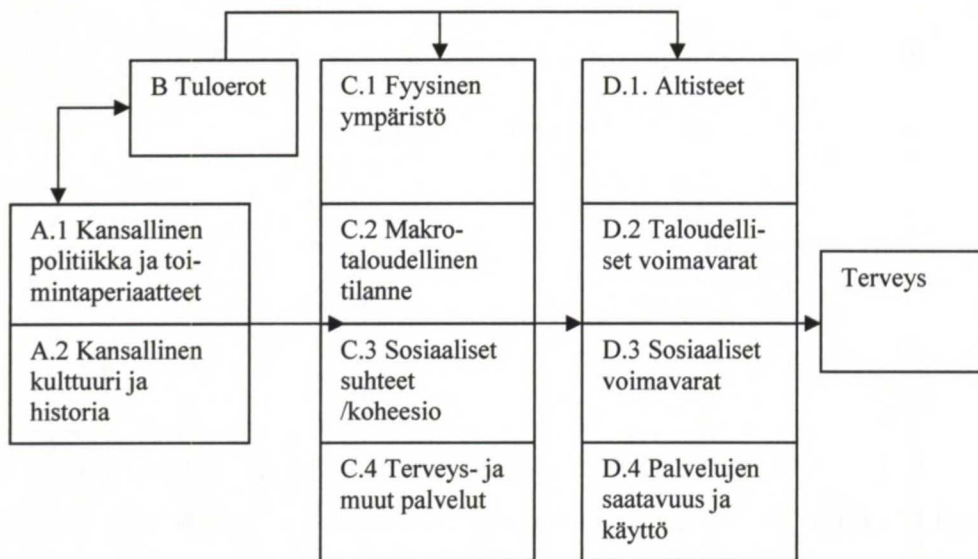
Lähde: Schmidt 2007

Yhteiskunnassa voi olla useita erilaisia vaikutusmekanismeja, joiden kautta tuloerot vahingoittavat terveyttä. Näiden mekanismien avulla pyritään selittämään erilaisia kausaalisia yhteyksiä tuloerojen ja terveyden välillä. Kawachi ja Kennedy (1999, 220–224) ovat tiivistäneet tuloerojen ja terveyden välistä yhteyttä selittävät tekijät kolmeen ryhmään, joita ovat vähäiset investoinnit inhimilliseen pääomaan, sosiaalisen pääoman heikkeneminen sekä stressiä lisäävä ja siten terveyttä vahingoittava ihmisten keskinäinen vertailu. Macinko ym. (2003, 415–419) tarkastelevat tuloerojen ja terveyden välistä yhteyttä jakamalla sen taustalla olevat tekijät kahteen pääryhmään: psykososiaalisiin ja neomateriaalisiin tekijöihin. Näitä tekijöitä havainnollistetaan kuviossa 7.

Kuviossa 7 yksilöiden terveyteen vaikuttavat tekijät on jaettu neljään pääryhmään (A-D) ja niiden alaryhmiin (1–4). Ensimmäinen ryhmä (ryhmä A) kuvaa tuloeroihin ja muihin terveyttä määrittäviin tekijöihin vaikuttavia kansallisia poliittisia, historiallisia ja kulttuuriin liittyviä tekijöitä. Tuloerot (ryhmä B) vaikuttavat puolestaan yhteiskunnan ja yksilön tasolla terveyttä määrittäviin tekijöihin. Tarkastelun oletuksena on, että tuloerojen yhteys terveyteen ei ole suora. Tuloerojen oletetaan vaikuttavan terveyteen välillisesti muiden tekijöiden kautta. Kolmannessa ryhmässä (ryhmä C) kuvataan terveyteen yhteiskunnan tasolla vaikuttavia tekijöitä. Näitä ovat esimerkiksi erilaiset ympäristötekijät, makrotaloudellinen tilanne, sosiaalinen koheesio sekä terveys- ja hyvinvointipalvelut. Yksilötasolla terveyteen vaikuttavia

tekijöitä (ryhmä D) ovat puolestaan erilaiset terveydelle haitalliset altisteet, yksilöiden taloudelliset ja sosiaaliset voimavarat sekä terveystaloudellisten palvelujen saatavuus ja niiden käyttö.

Kuvio 7. Tulorerojen ja terveyden välisen yhteyden käsitteellinen kehikko<sup>16</sup>



Lähde: Macinko ym. 2003, 410

Tuloerojen ja terveyden välisen yhteyden psykososiaalinen selitysmalli tarkastelee yksilöiden ja yhteisöjen psykososiaalisia ominaisuuksia, kuten ihmisten välistä luottamusta ja sosiaalista yhteenkuuluvuutta tuloerojen ja terveyden välisen yhteyden selittäjinä (Macinko ym. 2003, 415). Näitä tekijöitä ovat kuviossa 7 ryhmään C.3 merkityt sosiaaliset suhteet ja koheesio sekä ryhmään D.3 merkityt yksilöiden sosiaaliset voimavarat, eli niin sanottu sosiaalinen pääoma. Psykososiaalisen näkökulman mukaan yhteiskunnan suuret tuloerot vähentävät yksilöiden yhteenkuuluvuuden tunnetta, lisäävät epäluottamusta, heikentävät yksilöiden sosiaalisia voimavaroja ja yhteisön sosiaalista pääomaa (Macinko ym. 2003, 416). Sosiaaliset hierarkiat lisäävät yksilöiden keskinäistä vertailua ja siitä aiheutuvaa stressiä. Yhteisöjen sosiaalisen koheesion heikentyminen voi johtaa myös lisääntyneeseen rikollisuuteen. (ma. 416) Esimerkiksi tätä kautta suuret tuloerot voivat olla uhka myös yhteiskunnan varakkaimpien yksilöiden terveydelle ja hyvinvoinnille.

<sup>16</sup> Kuvion käsitteellinen kehikko perustuu Starfieldin ja Shin (1999) tutkimukseen sekä Macinkon ym. (2003, 436–445) tekemään katsaukseen tuloerojen ja terveyden välistä yhteyttä käsittelevästä kirjallisuudesta.



Tuloerojen ja terveyden välistä yhteyttä tarkastelevan neomateriaalisen selitysmallin mukaan aineelliset tekijät, kuten tulot ja elinolosuhteet ovat tärkeimpiä yksilöiden terveyteen vaikuttavia tekijöitä. Tuloerojen ja terveyden välisen yhteyden ajatellaan olevan seurausta yhteiskunnan vähäisistä investoinneista fyysiseen, inhimilliseen ja sosiaaliseen pääomaan, yksilöiden taloudellisten resurssien puutteesta sekä altistumisesta esimerkiksi huonoista elinolosuhteista johtuville vaaratekijöille. Vähäiset taloudelliset resurssit heikentävät yksilöiden mahdollisuuksia esimerkiksi sairauksien ehkäisyyn ja niiden hoitoon. Yhteiskunnan vähäiset investoinnit inhimilliseen ja sosiaalisiin pääomaan sekä fyysiseen ympäristöön heikentävät erityisesti huono-osaisten mahdollisuuksia esimerkiksi riittävän koulutuksen hankkimiseen ja hyviin elinolosuhteisiin. (Macinko ym. 2003, 418) Kuviossa 7 näitä taloudellisia tekijöitä havainnollistetaan ryhmissä C.2 ja D.2.

Kuviossa 7 terveydenhuollon vaikutus tuloerojen ja terveyden väliseen yhteyteen ilmenee kahdella tasolla: koko yhteiskunnassa terveystalouden rakenteen ja saatavuuden (ryhmä C.4) sekä yksilötasolla palveluiden käytön kautta (ryhmä D.4). Terveystalouden merkitystä tulojen ja terveyden välisen yhteyden selittäjinä on epäilty (ks. esim. Cutler ym. 2006, 26). Shi ja Starfield (2000, 546) osoittivat kuitenkin, että tuloeroilla ja perusterveydenhuollon lääkäreiden määrällä suhteessa väestömäärään oli tilastollisesti merkitsevä yhteys yksilöiden koettuun terveyteen Yhdysvalloissa. Osavaltioissa, joissa tuloerot olivat suurimmillaan, 16 prosenttia ihmisistä ilmoitti terveytensä kohtalaiseksi tai huonoksi, kun prosenttiosuus tasaisimman tulonjaon osavaltioissa oli 10. Myös perusterveydenhuollon lääkäreiden määrä suhteessa väestöön oli korkeampi niiden henkilöiden kohdalla, jotka ilmoittivat terveytensä hyväksi tai erinomaiseksi kuin terveytensä kohtalaiseksi tai huonoksi kokevien kohdalla. (ma. 546) Tutkijoiden mukaan erot perusterveydenhuollossa voivat olla eräs selittäjä, jonka kautta tuloerojen vaikutus välittyy haitallisesti terveyteen (ma. 552).

Edellä tuloerojen ja terveyden välistä yhteyttä selittäviä tekijöitä käsiteltäessä on oletettu, että huono terveys on seurausta alhaisesta tulotasosta ja, että tuloerot vaikuttavat haitallisesti terveyteen. Yksilöt eivät kuitenkaan ole välttämättä sairaita sen takia, että he ovat köyhiä (Macinko ym. 2003, 416). Alhainen tulotaso voi johtua myös huonosta terveydestä. Mikäli heikko terveydentila rajoittaa yksilön mahdollisuuksia tehdä työtä, heijastuu huono terveys tätä kautta tuloihin. Yksilöt voivat olla valikoituneet terveydentilansa perusteella eri tuloluokkiin. (ma. 416) Tällöin tuloerot olisivat osin seurausta terveyden eriarvoisesta jakautumisesta yksilöiden kesken, joten terveyseroja kaventamalla olisi mahdollista vähentää

myös tulojen eriarvoista jakautumista väestössä (Deaton 2003, 133). Terveysten vaikuttavat toimenpiteet heijastuisivat siten esimerkiksi parantuneen työkyvyn kautta tuloihin. Mikäli terveydentilaa parantamalla voitaisiin vaikuttaa yksilöiden työntekomahdollisuuksiin, olisi terveyseroja kaventamalla mahdollista supistaa tuloeroja. Tällöin terveyteen ja sen eriarvoiseen jakautumiseen vaikuttavat toimenpiteet heijastuisivat myös tuloeroihin.

## 4 TULOEROT JA TERVEYS: KATSAUS EMPIIRISEEN KIRJALLISUUTEEN

### 4.1 Yleistä

Tuloerojen ja terveyden välistä yhteyttä on tutkittu liki 30 vuoden ajan. Rodgers (1979) tutki tuloerojen ja väestön terveyden välistä yhteyttä 1970-luvulla ja havaitsi, että tuloeroilla oli yhteys elinajanodotteeseen ja lapsikuolleisuuteen: suuret tuloerot olivat yhteydessä korkeaan kuolleisuuteen ja tasaisemman tulonjaon maissa elinajanodote oli tulojen suhteen eriarvoisempia maita korkeampi. Sitten tuloerojen ja terveyden välistä yhteyttä selvittäneiden tutkimusten määrä on kasvanut voimakkaasti, eikä asiasta ole saavutettu yhteisymmärrystä. Wilkinson ja Pickett (2006) ovat laatineet kirjallisuuskatsauksen, jossa mainitaan kaikkiaan 155 julkaisua ja niiden 168 tuloerojen ja terveyden välistä yhteyttä selvittänyttä analyysia. Näistä analyyseista 70 prosenttia tuki täysin tai osittain esitettyä hypoteesia tuloerojen ja terveyden välisestä yhteydestä. Loput 30 prosenttia kumosivat sen. Seuraavassa luodaan katsaus aiheesta julkaistuihin tutkimuksiin ja niiden tuloksiin katsausartikkelien<sup>17</sup> pohjalta.

Tuloerojen ja terveyden välistä yhteyttä on tutkittu sekä maiden sisällä että niiden välillä. Näkökulmasta riippuen tutkimuksia on ryhmitelty kansainvälisiin väestötasoihin tutkimuksiin, maan sisäisiin aggregaattitason tutkimuksiin sekä maan sisäisiin monitasoihin tutkimuksiin. Viimeksi mainituissa tutkimuksissa on selvitetty tuloerojen yhteyttä yksilöiden terveyteen, kun taas maan sisäisissä aggregaattitason tutkimuksissa on tarkasteltu tuloerojen ja terveyden välistä yhteyttä aluetasolla. (ks. esim. Wagstaff & van Doorslaer 2000, Lynch ym. 2004) Tutkimukset, joissa on selvitetty tuloerojen yhteyttä väestön terveyteen eri maiden välillä, luokitellaan väestötasoihin tutkimuksiin (ks. esim. Rodgers 1979).

---

<sup>17</sup> Wagstaff & van Doorslaer (2000), Macinko ym. (2003), Lynch ym. (2004), Subramanian & Kawachi (2004), Wilkinson & Pickett (2006)



Tällä hetkellä tuloerojen ja terveyden välistä yhteyttä selvittävässä tutkimuksissa ollaan siirtymässä väestötasoisista tutkimuksista kohti yksilötasoisella aineistolla tehtyjä tutkimuksia (Wagstaff & van Doorslaer 2000, 552). On myös esitetty, että ihannetapauksessa tuloerojen ja terveyden välistä yhteyttä tulisi selvittää analysoimalla samanaikaisesti sekä yksilötason että yhteisötason muuttujien vaikutusta monitasoisten tilastollisten menetelmien avulla (ks. Dahl ym. 2006, 2563).

Subramanian ja Kawachin (2004) laatimassa katsauksessa käydään läpi 21 Yhdysvalloissa ja sen ulkopuolella monitasoisella aineistolla tehtyä tutkimusta. Tutkijat havaitsivat, että yhdeksässä Yhdysvalloissa tehdyssä tutkimuksessa tuloerojen ja terveyden välillä oli positiivinen yhteys. Ruotsissa, Tanskassa, Iso-Britanniassa, Japanissa, Uudessa-Seelannissa ja Chilessä tehdyissä tutkimuksissa tuloerojen ja terveyden välinen yhteys havaittiin ainoastaan viimeiseksi mainitussa maassa. Tuloeroilla ja terveydellä on esitetty olevan yhteys ainoastaan sellaisissa maissa, joissa eriarvoisuus on suurta, kuten esimerkiksi Yhdysvalloissa tai Chilessä (ks. Subramanian & Kawachi 2004, 81). Myös tuloerojen ja terveyden välistä yhteyttä tarkastelleessa kiinalaisessa tutkimuksessa havaittiin tuloerojen ja koetun terveydentilan välillä yhteys siten, että tuloerot olivat haitallisia terveydelle eriarvoisuuden ollessa suurta. Kiinassa tuloerot ovatkin kasvaneet voimakkaasti viime vuosikymmeninä. (ks. Li & Zhu 2006)

Maantieteelliset alueet, joiden perusteella alueiden tuloeroja on laskettu, ovat myös vaihdelleet eri tutkimuksissa. Wilkinson ja Pitkett (2006) mainitsevat kirjallisuuskatsauksessaan, että suurille maantieteellisille alueille laskettuja tuloeroja käyttäneissä tutkimuksissa on havaittu useammin tuloerojen ja terveyden välinen yhteys kuin tutkimuksissa, joissa tuloerot on laskettu pienille alueille. On esitetty, että maantieteellisesti pienillä alueilla tuloerot eivät heijasta riittävästi yhteiskunnallista eriarvoisuutta, eikä tuloerojen yhteyttä terveyteen tästä syystä esiinny (Subramanian & Kawachi 2004, 82). Myös tuloeromittarit ovat vaihdelleet eri tutkimuksissa (ks. esim. Macinko ym. 2003, 423–424).

Tutkimuksissa on käytetty myös useita erilaisia, terveyttä kuvaavia muuttujia (ks. Macinko ym. 2003, Subramanian & Kawachi 2004). Empiirisessä kirjallisuudessa havaittava suuntaus on, että sairastavuuden subjektiiviseen arviointiin perustuvia terveyden indikaattoreita käytetään yhä useammin terveyttä karkeasti kuvaavien mittareiden, kuten kuolleisuuden sijaan (Wagstaff & van Doorslaer 2000, 552).



Seuraavassa tarkastellaan yksityiskohtaisemmin muutamia tuloerojen ja terveyden välistä yhteyttä selvittänyttä tutkimusta. Esiteltävien tutkimusten valinnassa on käytetty kolmea kriteeriä. Aluksi tarkastellaan tutkimuksia, jotka ovat selvittäneet tuloerojen ja terveyden välistä yhteyttä Suomessa. Vertailtavuuden lisäämiseksi käsitellään myös Ruotsissa, Tanskassa ja Norjassa tehtyjä tutkimuksia, sillä Pohjoismaat ovat monessa suhteessa samankaltaisia. Pohjoismaissa esimerkiksi tuloerot ja suhteellinen köyhyys ovat varsin alhaisia, ja terveystaloudissa julkisten palveluntarjoajien ja verorahoituksen osuus on merkittävä (Kautto ym. 2004, 264, Rostgaard & Lehto 2004, 163). Kolmas valintakriteeri oli tutkimusten aineisto. Kaikissa esiteltävissä tutkimuksissa tarkastellaan alueiden tuloerojen ja yksilöiden terveyden välistä yhteyttä ja tutkimuksissa hyödynnetään monitasoisia aineistoja.

Tarkasteltavissa tutkimuksissa on pääsääntöisesti käytetty terveyden mittarina kuolleisuutta. Ainoastaan suomalaisessa Böckermanin ym. (2007) tutkimuksessa selitettävänä muuttujina käytettiin subjektiiviseen arviointiin perustuvia terveyden indikaattoreita.

#### 4.2 Tuloerot ja terveys Suomessa

Blomgren ym. (2004) ovat selvittäneet alueellisten tekijöiden yhteyttä työikäisten miesten alkoholikuolleisuuteen Suomessa. Tutkimuksen aineisto oli kaksitasoinen sisältäen tietoja sekä yksilöistä että alueista, joilla henkilöt asuivat. Tutkimus perustui vuoden 1990 väestötietoihin, jotka yhdistettiin vuosien 1991–1996 kuolintietoihin. Tutkimusaineisto kattoi 1,1 miljoonaa 25–64-vuotiaasta miestä 84 alueelta (NUTS 4)<sup>18</sup> Suomessa.

Tutkimuksessa alkoholin käyttöön liittyvää kuolleisuutta selitettiin yksilö- ja aluetason muuttujilla. Yksilötason muuttujia olivat ikä, koulutus, sosioekonominen asema, siviilisääty ja äidinkieli. Aluetason muuttujina olivat puolestaan Gini-kertoimella mitatut tuloerot vuonna 1990, alueen sosioekonominen rakenne<sup>19</sup>, kaupungistumisen aste ja ruotsinkielisen väestön osuus. Lisäksi aluetason selittävänä muuttujana oli sosiaalinen koheesio, jota mitattiin alueen äänestysprosentilla sekä yksinasuvien, eronneiden ja yksinhuoltajaperheiden väestöosuudella.

<sup>18</sup> NUTS-aluejako (Nomenclature des Unités Territoriales Statistiques) on Euroopan Unionin virallinen alueluokitusjärjestelmä. Suomessa NUTS 4-taso koostuu seutukunnista. (Niemi 2004, 6)

<sup>19</sup> Muuttujaan sisältyi ruumiillista työtä tekevien osuus sekä kotitalouksien OECD-kulutussyksikköä laskettu mediaanitulo alueella vuonna 1990 ja alueen keskimääräinen työttömyysaste vuosina 1990 ja 1993.

Aineisto analysoitiin tutkimuksessa monitasoisella Poisson-regressioanalyysillä (ks. monitasoisesta Poisson-regressioanalyysimenetelmästä esim. Blomgren 2005, 268–288).

Tutkimuksessa tuloeroilla ja alueen kotitalouksien mediaanitulolla ei havaittu tilastollisesti merkitsevää yhteyttä miesten alkoholikuolleisuuteen. Tutkijoiden mukaan tulos voi johtua Suomen suhteellisen tasaisesta tulonjaosta ja siitä, että tuloerot alueiden välillä ovat melko pienet. Tuloeroista ja kotitalouksien mediaanitulosta poiketen useilla muilla aluetason muuttujilla oli tutkimuksessa tilastollisesti merkitsevä yhteys miesten alkoholikuolleisuuteen. Tulosten mukaan alkoholikuolleisuus oli suurimmillaan niillä alueilla, joilla ruumiillista työtä tekevien työntekijöiden osuus sekä työttömyysaste olivat korkeita ja sosiaalinen koheesio alhainen. Tutkimus osoittaa, että alkoholikuolleisuuden alueittaisia eroja ei voida selittää pelkästään yksilötekijöissä esiintyvien alueellisten erojen perusteella.

Tutkimuksessa kaikilla yksilötason selittävillä muuttujilla oli tilastollisesti merkitsevä yhteys miesten alkoholin käyttöön liittyvään kuolleisuuteen. Alkoholikuolleisuus nousi iän myötä ja laski koulutustason ja sosioekonomisen aseman noustessa. Naimisissa olevien ja suomenruotsalaiseen väestöön kuuluvien miesten alkoholikuolleisuus oli muita miehiä pienempää.

Martikainen ym. (2004) ovat puolestaan selvittäneet alueellisten tekijöiden vaikutusta itsemurhakuolleisuuteen Suomessa. Tutkijoiden mukaan alkoholikuolleisuuden ohella itsemurhakuolleisuus on sopiva mittari aluetekijöiden terveysvaikutusten arvioimiseksi (Martikainen ym. 2004, 325). Tutkimuksessa käytetty yksilötason aineisto perustui vuoden 1990 väestötietoihin ja kattoi 15–99-vuotiaat suomalaiset. Tiedot yhdistettiin vuosien 1991–2001 kuolintietoihin. Aluetason aineisto kattoi yhteensä 85 aluetta (NUTS 4) Suomessa.

Tutkimuksessa alkoholin käyttöön liittyviä ja muita itsemurhakuolemia selitettiin yksilön sukupuolella, iällä, sosioekonomisella asemalla, kotitalouden tuloilla, asunnon omistuksella, taloudellisella toimeilla, siviilisäädellä, perhetyypillä, perheen koolla sekä äidinkielellä. Alueelliset selittävät muuttujat olivat Gini-kertoimella mitatut tuloerot vuonna 1990 sekä alueen sosioekonominen rakenne ja sosiaalinen koheesio. Kaksi jälkimmäistä muuttujaa määriteltiin kuten edellä (vrt. Blomgren ym. 2004). Tutkimusmenetelmä oli sama kuin edellä esitetyssä Blomgrenin ym. (2004) tutkimuksessa.



Martikaisen ym. (2004) tutkimuksen mukaan itsemurhakuolleisuus oli erityisen korkea alueilla, joilla ruumiillista työtä tekevien osuus sekä työttömyysaste olivat korkeita ja perheiden sosiaalinen koheesio ja äänestysprosentti alhaisia. Vaikka alueen sosioekonomisen rakenteen ja sosiaalisen koheesion vaikutus itsemurhakuolleisuuteen heikkeni jonkin verran yksilötekijät vakioitaessa, säilyi näiden tekijöiden vaikutus itsemurhakuolleisuuteen yksilötekijöiden kontrolloinnista huolimatta.

Tutkimuksen mukaan tuloeroilla ja alueen kotitalouksien mediaanitulolla ei ollut selvää ja johdonmukaista yhteyttä itsemurhakuolleisuuteen Suomessa. 15–64-vuotiaiden naisten kohdalla tuloeroilla havaittiin kuitenkin vähäinen yhteys naisten alkoholin käyttöön liittyviin itsemurhiin, kun yksilötason selittävistä muuttujista huomioitiin ainoastaan ikä. Itsemurhat olivat harvinaisempia aluilla, joissa tuloerot olivat suuret. Miesten osalta tuloeroilla ei löytynyt yhteyttä itsemurhakuolemiin, kun yksilötason selittävistä muuttujista huomioitiin ainoastaan ikä. Kun kaikki yksilötason ja aluetason muuttujat otettiin malliin mukaan, tuloeroilla ei ollut tilastollisesti merkitsevää yhteyttä miesten ja naisten itsemurhakuolleisuuteen. Myöskään alueen kotitalouksien mediaanitulon yhteys ei ollut tilastollisesti merkitsevä. Tulos on yhtenevä Blomgrenin ym. (2004) tutkimustulosten kanssa.

Böckerman ym. (2007) selvittivät edellä esitetyistä tutkimuksista poiketen maakuntien tuloerojen yhteyttä terveyteen Suomessa vuosina 1993–2005. Tutkimuksessa käytettiin aineistona Kansanterveyslaitoksen tekemää kyselytutkimusta Suomalaisen aikuisväestön terveyskäyttäytyminen ja terveys, johon yhdistettiin Tilastokeskuksen tulonjakotilastosta saatavat tiedot maakuntien tuloeroista. Tuloeroja mitattiin tutkimuksessa käytettävissä tuloista lasketulla Gini-kertoimella ja ekvivalenssiskaalana käytettiin OECD-skaalaa. Terveysmuuttujia oli yhteensä kuusi: hyvä itse arvioitu terveys, hyvä fyysinen terveys, hyvä psyykinen terveys, ei lääkkeiden käyttöä, ei sairaslomia ja ei työkyvyttömyyseläkkeellä. Kaikki muuttujat olivat indikaattorimuuttujia saaden arvoja 0 tai 1. Selittävinä muuttujina olivat tuloerojen ohella sukupuoli, ikä, siviilisääty, koulutusvuodet sekä alueen reaalisten käytettävissä olevien tulojen logaritmi. Estimoinnit suoritettiin probit-mallin avulla yhdistetyllä aineistolla, sekä naisille ja miehille erikseen.

Tutkimustulosten mukaan maakunnan tuloeroilla ei ollut kaiken kaikkiaan yhteyttä lisääntyneeseen sairastavuuteen Suomessa vuosina 1993–2005. Tuloksissa havaittiin tosin tuloerojen lievä lääkkeiden kulutusta nostava vaikutus, kun estimointi suoritettiin yhdistetyllä



aineistolla. Tulokset eivät olleet tilastollisesti merkitseviä tutkimuksessa käytettyjen viiden tai yhden prosentin luottamustasolla, eikä vaikutus ollut kovin merkittävä. Lääkkeiden kulutuksen todennäköisyys kasvoi 0,8 prosenttiyksikköä tuloerojen nousun myötä vuosina 1993–2005. Kun estimointi suoritettiin erikseen naisille ja miehille, havaittiin naisten kohdalla tuloerojen kasvun käänteinen yhteys hyvään fyysiseen terveyteen ja ei-työkyvyttömyyseläkkeellä oloon. Miesten kohdalla tuloerojen kasvulla ei ollut sairastavuutta lisäävää vaikutusta.

#### 4.3 Tuloterot ja terveys Ruotsissa, Tanskassa ja Norjassa

Gerdtham ja Johannesson (2004) ovat tutkineet tuloerojen ja kuolleisuuden yhteyttä Ruotsissa. Tutkijoiden tavoitteena oli testata tulojen eriarvoisuuden hypoteesin ohella myös absoluuttisen ja suhteellisen tulon hypoteesia, joten tutkimuksessa kuolleisuusriski estimoitin yksilön tulojen, kunnittain lasketun keskimääräisen tulon ja tuloerojen funktiona. Tutkimuksen aineisto perustui ruotsalaiseen elinolotutkimukseen, joka yhdistettiin tulo- ja kuolintietoja sisältäviin aineistoihin. Koko tutkimusaineisto kattoi tiedot yli 40 000 20–84-vuotiaasta ruotsalaisesta, joita haastateltiin vuosina 1980–1986 ja seurattiin tämän jälkeen aina vuoteen 1996 asti. Kunnat (n=284) muodostivat tutkimuksen viiteryhmän. Lisäksi käytettiin kahta muuta maantieteellisesti määriteltyä viiteryhmiä, joita olivat maakunnat (n=24) ja paikalliset työmarkkinat (engl. local labour markets, n=100).

Tutkimuksessa selitettävänä muuttujina olivat elossaoloaika, jota mitattiin vuosien määrällä haastattelupäivän ja kuolinpäivän välillä sekä elossaolostatus (elossa=1, kuollut=0) seurantajakson (31.12.1996) lopussa. Selitettävänä muuttujina olivat alueen keskimääräinen tulo, Gini-kertoimella mitatut tuloterot ja henkilön vuotuiset tulot, jotka muodostuivat kotitalouden aikuista kohti lasketuista käytettävissä olevista tuloista ja nettovarallisuudesta. Gini-kertoimen ohella käytettiin myös useita muita tuloeromittareita. Lisäksi kontrollimuuttujina käytettiin lukuisia yksilötason muuttujia (ks. Gerdtham ja Johannesson 2004, 235). Estimoinnissa käytettiin Coxin suhteellisten vaarojen mallia (engl. Cox proportional hazard model). Tutkimuksessa sovellettiin myös probit-mallia ja duraatiomalleja, jotka tuottivat samanlaisia tuloksia kuin Coxin malli.

Tutkimuksessa tuloerojen yhteys kuolleisuuteen ei ollut tilastollisesti merkitsevä. Tulos ei muuttunut, vaikka tuloterot määriteltiin maantieteellisesti eri alueille useiden eri

tuloeromittareiden perusteella. Gertdhamin ja Johannessonin (2004) tutkimuksen mukaan tuloeroilla ja kuolleisuudella ei ole yhteyttä Ruotsissa. Tutkimuksen tulokset eivät tukeneet myöskään hypoteesia siitä, että tuloerot olisivat haitallisia ainoastaan yhteiskunnan huono-osaisten terveydelle.

Kuolleisuusriski laski tutkimuksen mukaan kuitenkin merkittävästi yksilön tulojen noustessa. Tulos tukee siten absoluuttisen tulon hypoteesia. Tulojen ja kuolleisuuden välinen yhteys osoittautui epälineaariseksi, eli tulojen kuolleisuusriskiä pienentävä vaikutus heikkeni yksilön tulotason noustessa. Kunnan keskimääräisen tulon yhteys kuolleisuuteen ei ollut tutkimuksessa tilastollisesti merkitsevä. Tulos muuttui kuitenkin, kun keskimääräinen tulo laskettiin kuntien sijaan maakunnittain. Tällöin tulokseksi saatiin, että kuolleisuusriski pieneni keskimääräisen tulon noustessa. Suhteellisen tulon hypoteesin mukaan kuolleisuuden tulisi kuitenkin nousta alueen keskimääräisen tulon kasvaessa. Gertdhamin ja Johannessonin (2004) tutkimustulokset eivät siten tukeneet suhteellisen tulon hypoteesia.

Myös Henriksson ym. (2006) ovat tutkineet tuloerojen ja kuolleisuuden välistä yhteyttä Ruotsissa. Tutkimuksen perusjoukko muodostettiin vuoden 1990 väestötiedoista ja se kattoi 2,57 miljoonaa 40–64-vuotiasta henkilöä. Perusjoukon tiedot yhdistettiin vuosien 1991–1998 kuolintietoihin. Tutkimushenkilöiden tulot ilmoitettiin kotitalouden kulutusyksikköä kohden laskettuna käytettävissä olevana tulona. Tulojen lisäksi yksilötason selittävinä muuttujina olivat ikä, sukupuoli ja sosioekonominen asema. Tuloerot laskettiin alueitten Gini-kertoimen sekä korkeimman (p90) ja matalimman (p10) tulotason suhdeluvun  $p90/p10$  avulla. Gertdhamin ja Johannessonin (2004) tavoin Henriksson ym. (2006) laskivat tuloerot kunnittain ( $n=284$ ). Tutkimuksessa hyödynnettiin kahta eri menetelmää: sekä tavallista yksitasoista että monitasoista Poisson-regressioanalyysia.

Ensimmäisellä menetelmällä saatujen tulosten perusteella tuloeroilla oli positiivinen ja tilastollisesti merkitsevä yhteys kuolleisuuteen, kun selittävinä muuttujina olivat tuloerojen lisäksi vain yksilön ikä ja sukupuoli. Kun malliin lisättiin loput muuttujat, riskisuhde kasvoi entisestään osoittaen yhä merkittävämpää tuloerojen ja kuolleisuuden välistä yhteyttä. Myös tuloilla ja kuolleisuudella havaittiin tällöin selvä positiivinen yhteys.

Jälkimmäisellä menetelmällä saadut tulokset poikkesivat selvästi ensimmäisellä menetelmällä saaduista tuloksista. Kun kuolleisuutta selitettiin tuloeroilla, iällä ja sukupuolella, tulonjaon ja



kuolleisuuden välinen yhteys osoittautui edellisiin tuloksiin nähden päinvastaiseksi: tuloerojen kasvaessa riskisuhde pieneni. Kun malliin lisättiin loput yksilötason muuttujat, riskisuhde kasvoi eikä tuloerojen ja kuolleisuuden välinen yhteys ollut enää tilastollisesti merkitsevä. Tulojen ja sosioekonomisen aseman yhteys kuolleisuuteen oli kuitenkin edelleen positiivinen.

Henrikssonin ym. (2006) mukaan monitasoinen regressioanalyysi on sopivin tapa analysoida edellä kuvattua aineistoa. Näin ollen tutkijoiden mukaan tuloeroilla ja kuolleisuudella ei ollut saatujen tulosten perusteella yhteyttä Ruotsissa. Tämä tulos tuloerojen ja terveyden välisestä yhteydestä on siten yhdenmukainen Gerdthamin ja Johannessonin (2004) sekä Blomgrenin ym. (2004) ja Martikaisen ym. (2004) Suomea koskevien tutkimustulosten kanssa.

Osler ym. (2002) selvittivät tuloerojen ja kuolleisuuden välistä yhteyttä Tanskassa. Tutkimuksen aineistona käytettiin kahta Kööpenhaminassa tehtyä pitkittäistä väestötutkimusta. Tiedot tutkimusten kohdehenkilöiden kuolleisuudesta, asumisesta, tuloista ja koulutuksesta saatiin muista tanskalaisista rekisterilähteistä. Lopullinen tutkimusaineisto käsitti yhteensä 25 728 henkilöä, joista naisia oli hieman yli puolet. Kohdehenkilöiden kuolleisuutta seurattiin vuosina 1980–1999 keskimäärin 12,8 vuoden ajan.

Tutkimuksessa yksilön kuolleisuutta selitettiin kotitalouden bruttotuloilla ja alueittain lasketuilla tuloeroilla. Osler ym. (2002) laskivat tuloerot jokaiselle seurakunta-alueelle (engl. parish, n=149) Kööpenhaminassa. Tuloeromittarina oli mediaanitulo-osuus, joka saatiin laskemalla kotitalouksien köyhimmän 50 prosentin bruttotulojen osuus kaikista kotitalouksien tuloista. Lisäksi selittävinä muuttujina käytettiin yksilön painoindeksiä, tupakointia, alkoholin käyttöä, vapaa-ajan fyysistä aktiivisuutta, koulutusta ja kotitalouden rakennetta. Alueitasoisina muuttujina käytettiin myös keskimääräisiä kotitalouden tuloja sekä alueen lapsiperheiden osuutta kaikista kotitalouksista. Estimoinnissa sovellettiin Coxin suhteellisten vaarojen mallia.

Tutkimustulosten mukaan tuloerojen ja kuolleisuuteen välillä ei ollut yhteyttä Tanskassa. Naisten kohdalla tuloeroilla ja kuolleisuudella ei ollut yhteyttä, vaikka tulojen lisäksi edellä mainittuja selittäviä muuttujia ei otettu estimoidussa mallissa lainkaan huomioon. Miesten osalta tuloerojen ja kuolleisuuden välinen yhteys hävisi, kun kyseessä olevat tekijät otettiin malliin mukaan. Sekä miehillä että naisilla tulot ja kuolleisuus olivat kuitenkin yhteydessä toisiinsa siten, että suurituloisilla kuolleisuus oli alhaisin.

Norjassa tuloerojen ja terveyden välistä yhteyttä ovat tutkineet Dahl ym. (2006). Tutkijoiden ensimmäisenä tavoitteena oli selvittää, miten tulonjaon alueelliset erot olivat yhteydessä kuolleisuuteen Norjassa 1990-luvulla. Toisena tavoitteena oli selvittää, esiintyikö tuloerojen terveydelle haitallisissa vaikutuksissa eroja erilaisessa yhteiskunnallisessa asemassa olevien yksilöiden välillä. Tutkimuksessa testattiin siten sekä tulojen eriarvoisuuden hypoteesin vahvaa että heikkoa versiota.

Tutkimuksen aineisto perustui Norjan Tilastokeskuksen erilaisista hallinnollisista rekistereistä kokoamaan FD-Trygd-tietokantaan, joka sisältää tietoja kaikista vuoden 1993 alussa Norjan kansalaisina olleista henkilöistä. Tietokannan tiedot yhdistettiin tutkimuksessa erilaisiin muihin hallinnollisiin rekistereihin ja väestötilastoihin aina vuoteen 1999 asti. Koko tutkimusaineisto kattoi yhteensä lähes 2,2 miljoonaa vuosina 1927–1968 syntynyttä norjalaista miestä ja naista.

Tutkimuksessa selitettävä muuttuja oli vuosina 1994–1999 havainnoitu kuolleisuus (esim. henkilö kuollut vuosina 1994–1999 tai elossa tarkastelujakson lopussa 1999). Selittäviä muuttujia olivat muun muassa alueiden keskimääräinen tulo ja tuloerot. Alueluokituksena käytettiin Euroopan unionin NUTS-aluejaon mukaista alueluokitusjärjestelmää. Yhteensä alueita oli 88 ja ne muodostivat NUTS 4-tason. Samaa aluejakoa käytettiin myös edellä esitetyissä Suomea koskeneissa Blomgrenin ym. (2004) ja Martikaisen ym. (2004) tutkimuksissa. Alueellisia tuloeroja Dahlin ym. (2006) tutkimuksessa mitattiin Gini-kertoimella. Aluemuuttujat laskettiin kotitalouden kulutusyksikköä kohden laskettujen käytettävissä olevien tulojen perusteella. Muita selittäviä muuttujia olivat yksilön ikä, sukupuoli, siviilisääty, koulutustaso, tulot ja saadut terveyteen liittyvät sosiaaliuudet. Estimoinnissa sovellettiin monitasoista logistista regressioanalyysia (ks. menetelmästä esim. Blomgren 2005, 282–286).

Tulosten mukaan alueelliset tuloerot olivat yhteydessä kokonaiskuolleisuuteen Norjassa 1990-luvulla. Tutkimus osoittaa, että kuolleisuus lisääntyi alueellisten tuloerojen kasvaessa. Vuosina 1994–1999 kuolleisuus oli noin 25–30 prosenttia korkeampi alueilla, joissa tuloerot olivat suurimmat verrattuna tasaisimman tulonjaon alueisiin. Tuloerojen ja kuolleisuuden välillä säilyi tilastollisesti merkitsevä positiivinen yhteys, vaikka edellä mainitut yksilötason tekijät otettiin malliin mukaan. Dahl ym. (2006) havaitsivat myös, että tuloerot vaikuttivat kaikkein haitallisimmin sosioekonomiselta asemaltaan huono-osaisiin. Sosioekonomista



asemaa mitattiin tutkimuksessa tuloilla, koulutuksella ja terveyteen liittyvillä sosiaalietauksilla. Dahlin ym. (2006) tutkimus tukee siten sekä tulojen eriarvoisuuden hypoteesin vahvaa että heikkoa versiota.

Tuloeroilla ja terveydellä on arveltu olevan yhteys ainoastaan maissa, joissa eriarvoisuus on suurta. Suomen, Ruotsin ja Tanskan osalta edellä esitetyt tutkimustulokset tukevat tätä havaintoa. Pohjoismaissa tuloerot ovat varsin pienet, eikä tuloeroilla ja terveydellä ole havaittu yhteyttä näissä tutkimuksissa. Norjaa koskeva Dahlin ym. (2006) tutkimustulos poikkeaa kuitenkin edellä mainituista pohjoismaalaisista tutkimuksista. Dahl ym. (2006, 2570) esittävät, että tasaisen tulonjaon maissa tuloerojen ja terveyden välistä yhteyttä tulisi selvittää tarkemmin. Tutkijoiden mukaan eräs lisätarkastelukohde on tutkimuksessa käytettävä aluejako. Maantieteellisesti pienille alueille lasketut tuloerot eivät välttämättä heijasta riittävästi yhteiskunnallista eriarvoisuutta, eikä tuloerojen yhteyttä terveyteen näin ollen esiinny (Subramanian & Kawachi 2004, 82). Osler ym. (2002), Gertdham ja Johannesson (2004) sekä Henriksson ym. (2006) käyttivät tutkimuksissaan kuntia tai niitä pienempiä alueita aluejaon perustana, eivätkä havainneet tuloerojen ja kuolleisuuden välillä positiivista yhteyttä. Dahlin ym. (2006) tutkimuksessa tuloerot laskettiin maantieteellisesti edellä mainittuja suuremmille alueille, ja tutkimuksessa tuloerojen ja terveyden välillä oli positiivinen yhteys. Suomea koskevissa Blomgrenin ym. (2004) ja Martikaisen ym. (2004) tutkimuksissa aluejako oli sama kuin norjalaisessa Dahlin ym. (2006) tutkimuksessa, mutta Suomessa tuloeroilla ja kuolleisuudella ei havaittu tutkimuksissa yhteyttä. Dahl ym. (2007, 2571) arvelevat tuloksen olevan seurausta tutkimuksissa käytettyjen terveystuuttujien eroista tai alueiden välisten tuloerojen pienemmästä vaihtelusta Suomessa Norjaan verrattuna.

## 5 AINEISTO, MUUTTUJAT JA TUTKIMUSMENETELMÄT

Käsillä olevan tutkielman empiirisessä osiossa selvitetään, minkälainen yhteys maakuntien tuloeroilla on suomalaisten pitkäaikaissairastavuuteen ja itse arvioituun yleiseen terveydentilaan. Lisäksi tarkastellaan seutukuntien tuloerojen ja yksilöiden terveyden välistä yhteyttä. Tutkielmassa testataan empiirisesti sekä tulojen eriarvoisuuden hypoteesin vahvaa että heikkoa versiota. Seuraavassa käsitellään seikkaperäisesti tutkielman aineistoa, muuttujia ja tutkimusmenetelmiä.

### 5.1 Aineisto

Tutkielman aineiston muodostavat Tilastokeskuksen vuosien 2003 ja 2004 tulonjakotilastot (Income Distribution Statistics, IDS) sekä tulonjakotilaston yhteydessä edellä mainittuina vuosina Euroopan Unionin tilastoviraston Eurostatin tulo- ja elinolotutkimusta (EU-SILC) varten kerätyt yksilöiden terveystiedot, jotka yhdistetään tulonjakotilastoista saataviin tietoihin.

Tilastokeskuksen tulonjakotilasto on julkaistu vuosittain aina vuodesta 1977 lähtien vuosia 1981 ja 1985 lukuun ottamatta. Se on otantaan perustuva tutkimus, jonka lopullinen otos muodostuu noin 10 000 yksityiskotitaloudesta Suomessa. Vuodesta 2003 lähtien tulonjakotilaston otokseen poimitut kohdehenkilöt ovat olleet iältään vähintään 16-vuotiaita. Aiemmin tulonjakotilaston ikäraja oli 15 vuotta. Kohdehenkilöiden kanssa samassa asutokunnassa asuvat henkilöt yhdistetään kohdehenkilöihin kotipaikkatunnuksen perusteella (Tilastokeskus 2006c, 29). Tulonjakotilasto perustuu kiertävään paneeliotokseen, jossa sama kotitalous on mukana kahtena peräkkäisenä vuotena (mt. 29). Tässä tutkielmassa käytettävien tulonjakotilastojen lopullinen otoskoko oli 11 200 kotitaloutta vuonna 2003 ja 11 229 kotitaloutta vuonna 2004. Yksilöitä kotitalouksissa oli yhteensä 29 070 vuonna 2003 ja 29 112 vuonna 2004. Vuoden 2003 tulonjakotilastossa ensimmäistä kertaa tutkimuksessa mukana olleita kotitalouksia oli yhteensä 5 873, joista vuotta myöhemmin oli mukana 5 797.

Tulonjakotilasto sisältää tietoja kotitalouksien ja ammatissa toimivien henkilöiden tulojen määrästä ja niiden muutoksista sekä tulojen rakenteesta ja jakautumisesta (Tilastokeskus 2006c, 9). Tulonjakotilastossa käytettäviä tulokäsitteitä tarkastellaan luvussa 5.2. Erilaisten taloudellisten tietojen lisäksi tulonjakotilastossa on saatavilla tietoja esimerkiksi



kotitalouksien koosta ja asuinpaikasta sekä henkilöiden iästä, sukupuolesta, siviilisäädystä, koulutusasteesta ja sosioekonomisesta asemasta. Tulonjakotilaston sisältämät tiedot kerätään yhdistämällä erilaisista hallinnollisista rekistereistä koottuja ja kotitalouksilta haastatteluiden avulla saatuja tietoja.

Vuodesta 2003 lähtien tulonjakotilaston tietojen keruun yhteydessä on kerätty tietoja eurooppalaista tulo- ja elinolotutkimusta (EU-SILC) varten. Eurooppalaista tulo- ja elinolotutkimusta varten kerätty aineisto (tässä: SILC-aineisto) sisältää haastatteluiden avulla kerättyjä tietoja esimerkiksi kotitalouksien sosiaalisesta syrjäytymisestä, asumisesta, kotitalouden jäsenten koulutuksessa olosta ja työmarkkina-asemasta sekä kohdehenkilöiden terveydestä ja terveyspalvelujen saannista. Tässä tutkielmassa hyödynnetään SILC-aineistosta saatavia kohdehenkilöiden terveystietoja vuosilta 2003 ja 2004. Nämä tiedot yhdistetään vuosien 2003 ja 2004 tulonjakotilastoista saatuihin taustatietoihin kunkin henkilön kotitaloudelle annetun numeron perusteella<sup>20</sup>. Tulonjakotilastossa taustatietoja on saatavilla kaikille kotitalouden jäsenille, mutta SILC-aineistossa terveystietoja on kotitalouden kohdehenkilöille. Tämän vuoksi yhdistetty aineisto supistettiin kohdehenkilötasoiseksi. Saatu aineisto vastaa havaintomäärältään vuosien 2003 ja 2004 tulonjakotilastojen kotitalouksien lukumääriä.<sup>21</sup>

## 5.2 Tulokäsité ja ekvivalenssiskaalat

Kotitalouden käytettävissä olevat tulot on Tilastokeskuksen tulonjakotilaston keskeisin tulokäsité, joka toimii myös tutkielman empiirisen osion tulokäsittelenä. Tulonjakotilaston tulokäsitteitä tarkasteltiin aiemmin tutkielman luvussa 2.1.1. Edellä mainittu kokonaistulokäsité ei sisälly tutkielman empiiriseen osioon, joten henkilöiden käyttämien ilmaisten tai subventoitujen palveluiden arvo jää tarkastelun ulkopuolelle.

Tutkielmassa tarkastelukohteena ovat kotitalouksissa asuvat yksilöt. Kotitalouksien tulojen vertailu voi antaa harhaanjohtavan kuvan yksittäisen kotitalouden jäsenen hyvinvoinnista, koska kotitaloudet eroavat sekä kooltaan että rakenteeltaan toisistaan. Jotta yksilöt saadaan

<sup>20</sup> Tulonjakotilaston ja SILC-aineiston yhdistäminen suoritettiin SAS 9.1-ohjelmistolla.

<sup>21</sup> Vuonna 2003 tapahtuneen ikärajamuutoksen ja tulonjakotilaston kiertävän paneeliotoksen vuoksi kyseisen vuoden aineistossa noin puolet henkilöistä valittiin otantaan 15-vuoden vähimmäisiän perusteella vuonna 2002. Vuonna 2003 tutkimuksessa ensimmäistä kertaa mukana olleiden henkilöiden kohdalla otannassa käytettiin uutta 16-vuoden ikärajaa.

keskenään paremmin vertailukelpoisiksi, muunnetaan kotitalouden käytettävissä olevat tulot henkilömääräisiksi kotitalouden jäsenmäärän (per capita) sekä OECD-skaalan ja modifioidun OECD-skaalan avulla.

Tarkasteltaessa kotitalouden jäsentä kohti laskettuja käytettävissä olevia tuloja jaetaan kotitalouden käytettävissä olevat tulot kotitalouden jäsenmäärällä. OECD-skaalassa ja modifioidussa OECD-skaalassa kotitalouden jäsenet saavat eri painoja. OECD-skaalassa kotitalouden ensimmäinen aikuinen saa painon 1, muut aikuiset painon 0,7 ja 0–17-vuotiaat lapset painon 0,5 (Tilastokeskus 2006c, 36–37). Modifioidussa OECD-skaalassa kotitalouden ensimmäinen aikuinen saa puolestaan painon 1, muut 14 vuotta täyttäneet aikuiset painon 0,5 ja 0–13-vuotiaat lapset painon 0,3 (mt. 37). Esimerkiksi kahden aikuisen ja yhden, korkeintaan 13-vuotiaan lapsen kotitalouden yksikköluvuksi saadaan OECD-skaalalla  $1+0,7+0,5=2,2$  ja modifioidulla OECD-skaalalla  $1+0,5+0,3=1,8$ . Per capita-tarkastelussa yksikköluku olisi tässä tapauksessa luonnollisesti kolme. Modifioidulla OECD-skaalalla yhteiskulutushyödyt ovat suuremmat kuin OECD-skaalalla tai jäsenmäärällä laskettuina.

### 5.3 Muuttujien kuvaus

#### 5.3.1 Selitettävät muuttujat

SILC-aineistosta saadaan tietoja kohdehenkilöiden yleisestä terveydentilasta ja sairastavuudesta. Tutkielman selitettäviä muuttujia ovat henkilöiden pitkäaikaissairastavuus ja yleinen terveydentila. Taulukossa 1 esitetään kohdehenkilöiltä kysytyt kysymykset, joiden pohjalta selitettävät muuttujat on muodostettu. Selitettävien muuttujien kuvaukset esitetään myös liitteen 2 taulukossa 1. Koska SILC-aineistosta puuttui terveystietoja muutamilta kohdehenkilöiltä, jouduttiin aineistoa supistamaan jonkin verran sen alkuperäisestä koosta. Selitettävien muuttujien jakaumat vuosien 2003 ja 2004 otoksessa esitetään taulukoissa 2 ja 3. Muuttujien kuvauksellinen analyysi raportoidaan liitteen 2 taulukoissa 2 ja 3.



Taulukko 1. Selitettävien muuttujien kuvaus

Muuttujan nimi	Selite	Arvot
Pitkäaikaissairastavuus	Onko Teillä jokin pitkäaikainen sairaus, vaiva tai vamma? Pitkäaikainen: - jatkuva seuranta tai hoitoa vaativa - kestänyt tai sen odotetaan kestävän ainakin puoli vuotta	1 jos henkilöllä on pitkäaikainen sairaus, vaiva tai vamma eli henkilöllä on pitkäaikaissairauksia, 0 jos ei.
Yleinen terveydentila	Onko terveydentilanne nykyisin mielestänne hyvä, melko hyvä, keskitasoinen, melko huono, huono?	3 jos henkilön yleinen terveydentila on hyvä, 2 jos melko hyvä, 1 jos keskitasoinen, 0 jos melko huono tai huono.

Lähde: Tilastokeskus 2007

Taulukko 2. Pitkäaikaissairastavuus-muuttujan jakauma tutkimusaineistossa

Muuttuja	2003		2004	
	Havaintojen lukumäärä	%-osuus	Havaintojen lukumäärä	%-osuus
Pitkäaikaissairastavuus = 0	6575	60,72	6405	58,70
Pitkäaikaissairastavuus = 1	4254	39,28	4506	41,30
Yhteensä	10829	100,00	10911	100,00

Taulukko 3. Yleinen terveydentila-muuttujan jakauma tutkimusaineistossa

Muuttuja	2003		2004	
	Havaintojen lkm	%-osuus	Havaintojen lkm	%-osuus
Yleinen terveydentila = 0	957 <sup>a</sup>	8,84	996 <sup>b</sup>	9,13
Yleinen terveydentila = 1	2310	21,34	2273	20,85
Yleinen terveydentila = 2	2713	25,06	2796	25,64
Yleinen terveydentila = 3	4846	44,76	4839	44,38
Yhteensä	10826	100,00	10904	100,00

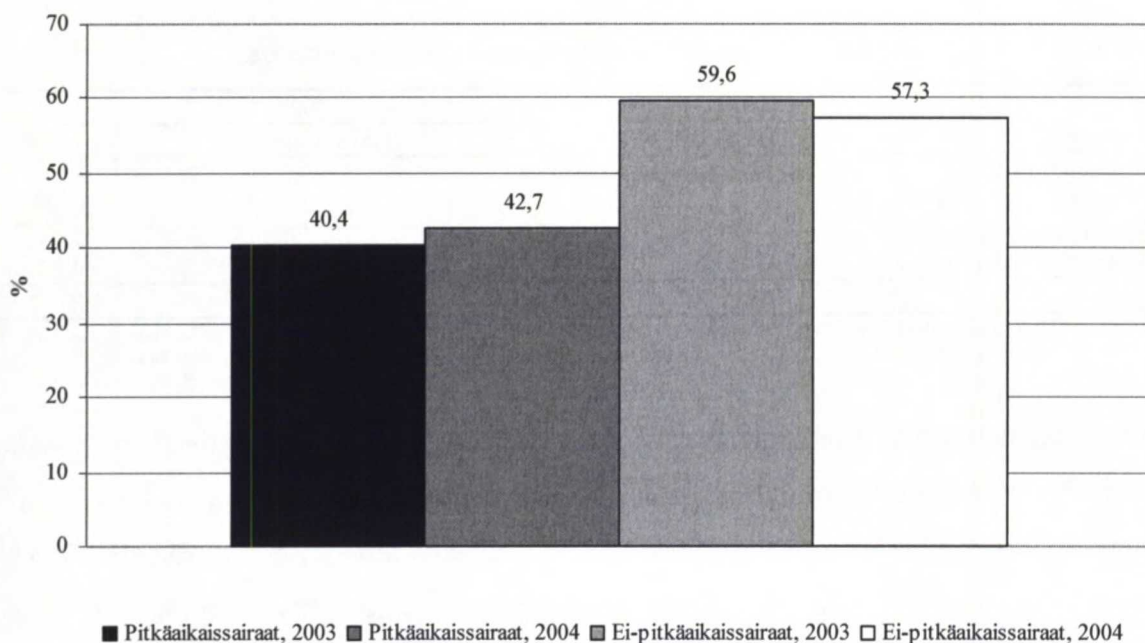
<sup>a</sup> Sisältää 244 henkilöä, joiden yleinen terveydentila oli huono ja 713 henkilöä, joiden yleinen terveydentila oli melko huono.<sup>b</sup> Sisältää 266 henkilöä, joiden yleinen terveydentila oli huono ja 730 henkilöä, joiden yleinen terveydentila oli melko huono.

Lisäksi selitettävänä muuttujana käytetään kunkin yleisen terveydentilan luokkaan kohdistettuja 15D-terveysmittarin luokkakeskiarvoja (ks. liite 9, taulukko 1). 15D-terveysmittari on yli 15-vuotiaalle aikuisväestölle kehitetty, terveyteen liittyvän elämänlaadun mittari (Sintonen 2001, 328 ja 334). Mittari selvittää terveyteen liittyvää fyysistä, psyykkistä ja sosiaalista elämänlaatua 15 ulottuvuuden osalta. Kyselyn avulla selvittävät 15 eri elämänlaadun osa-aluetta ovat hengitys, puhekyky, näkökyky, liikuntakyky, työkyky, energisyys, kuulo, syöminen, pidätyskyky, nukkuminen, sukupuolielämä, henkinen toimintakyky, ahdistuneisuus, masentuneisuus sekä kipu ja särky. Jokaista osa-aluetta koskeva kysymys sisältää viisi vastausvaihtoehtoa, joiden perusteella lasketaan yksi terveyteen liittyvää elämänlaatua kuvaava luku. 15D-mittari saa arvoja nollasta yhteen siten,

että arvon ollessa 1 henkilöllä ei ole ongelmia millään osa-alueella ja arvo 0 merkitsee kuolemaa. (ma. 329)

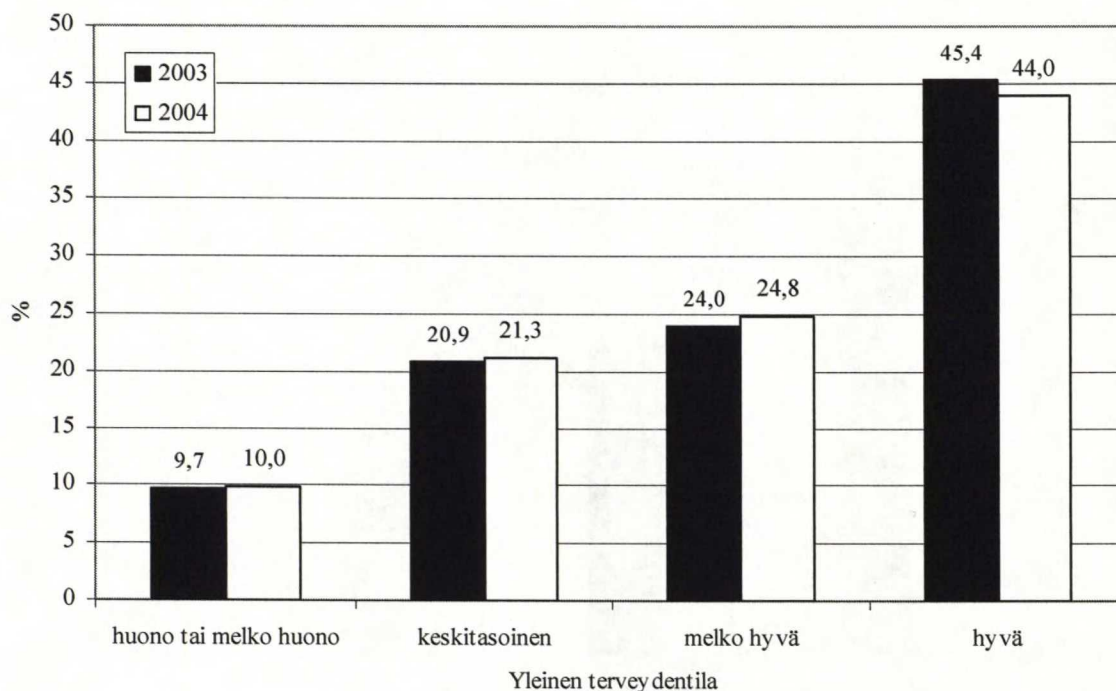
Seuraavassa tarkastellaan pitkäaikaissairastavuuden ja yleisen terveydentilan jakautumista yli 15-vuotiaiden suomalaisten osalta vuosina 2003 ja 2004. SILC-aineiston otoksen kohdehenkilöiden terveystiedot on korotettu korotuskertoimien avulla edustamaan perusjoukon tasoa. Kuvio 8 havainnollistaa pitkäaikaissairaiden ja ei-pitkäaikaissairaiden osuuksia Suomessa vuosina 2003 ja 2004. Kuviosta voidaan havaita, että pitkäaikaissairastavuus lisääntyi Suomessa vuodesta 2003 vuoteen 2004. Pitkäaikaissairaiden osuus kasvoi tällöin 2,3 prosenttiyksikköä. Vuonna 2004 jotain pitkäaikaista sairautta, vaivaa tai vammaa sairasti melkein 43 prosenttia yli 15-vuotiaista suomalaisista.

Kuvio 8. Pitkäaikaissairaiden ja ei-pitkäaikaissairaiden väestöosuudet Suomessa vuosina 2003 ja 2004, % yli 15-vuotiaista henkilöistä





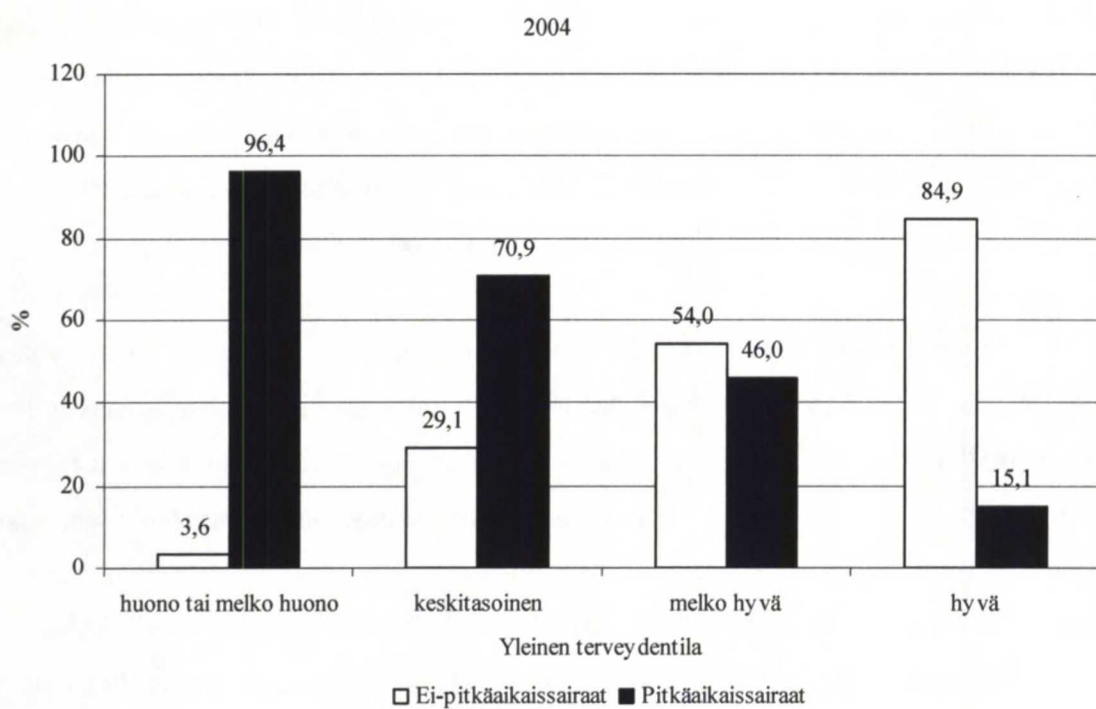
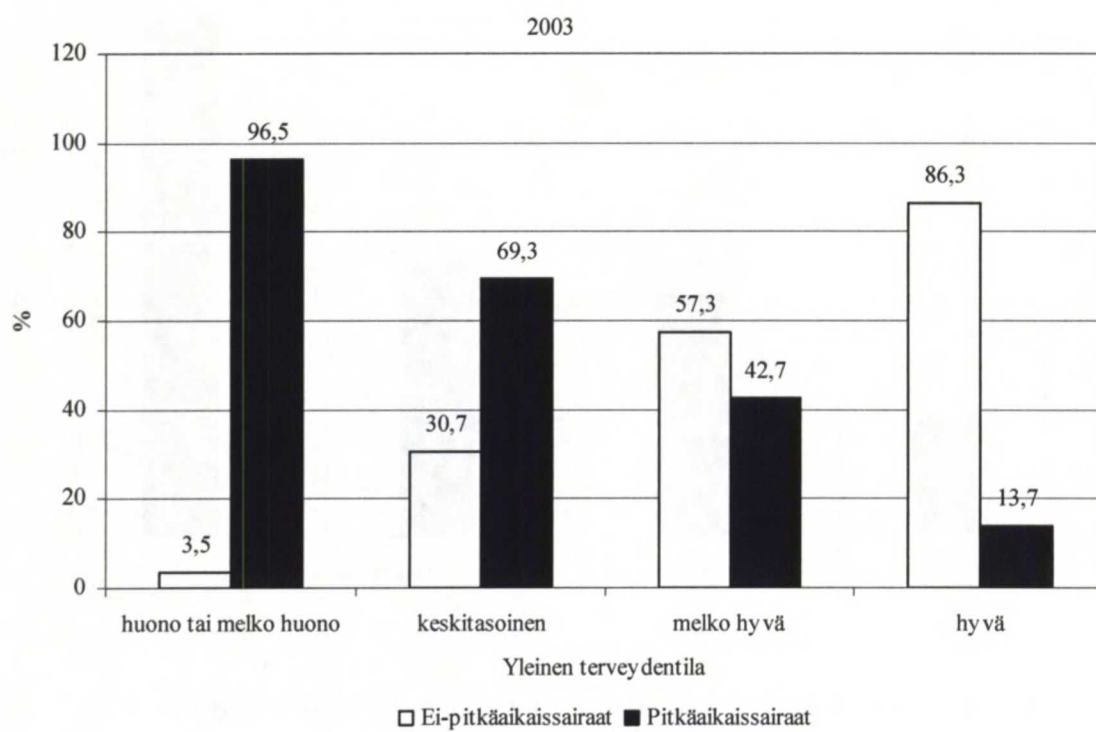
Kuvio 9. Yleisen terveydentilan jakautuminen Suomessa vuosina 2003 ja 2004, % yli 15-vuotiaista henkilöistä



Kuviossa 9 esitetään yleisen terveydentilan jakautuminen Suomessa vuosina 2003 ja 2004. Vuonna 2003 huonoksi tai melko huonoksi terveydentilansa koki 9,7 prosenttia yli 15-vuotiaista suomalaisista. Vuotta myöhemmin osuus oli kasvanut 0,3 prosenttiyksikköä. Myös melko hyväksi tai keskitasoiseksi terveydentilansa kokevien osuus kasvoi vuodesta 2003 vuoteen 2004. Vastaavasti terveydentilansa hyväksi kokevien osuus laski 1,4 prosenttiyksikköä vuodesta 2003 vuoteen 2004. Tästä huolimatta yli 40 prosenttia yli 15-vuotiaista suomalaisista koki terveydentilansa edelleen hyväksi vuonna 2004.

Kuvio 10 havainnollistaa yleisen terveydentilan ja pitkäaikaissairastavuuden välistä yhteyttä vuosina 2003 ja 2004. Kuviosta voidaan tehdä selkeä havainto pitkäaikaissairastavuuden ja yleisen terveydentilan välisestä yhteydestä. Huonoksi tai melko huonoksi terveytensä kokevien joukossa hieman yli 96 prosenttia ilmoitti sairastavansa pitkäaikaissairauksia. Hyväksi terveydentilansa kokevien joukossa pitkäaikaissairauksia sairastavien osuus oli selvästi alle 20 prosenttia vuosina 2003 ja 2004. Pitkäaikaissairaiden osuus laski tasaisesti sen mukaan, oliko tarkastelun kohteena huono tai melko huono, keskitasoinen, melko hyvä vai hyvä yleinen terveydentila.

Kuvio 10. Pitkäaikaissairaiden ja ei-pitkäaikaissairaiden osuudet yleisen terveydentilan eri luokissa vuosina 2003 ja 2004





### 5.3.2 Aluepiirteitä kuvaavat selittävät muuttujat

Tutkielman aluepiirteitä kuvaavissa muuttujissa aluejaon perustana on Euroopan Unionin virallinen alueluokitusjärjestelmä NUTS (Nomenclature des Unités Territoriales Statistiques). Alueelliset muuttujat perustuvat NUTS 3- ja NUTS 4-tasoihin. Suomessa NUTS 3-taso koostuu maakunnista. Ahvenanmaan maakuntaa koskevat luvut yhdistetään tutkielmassa Varsinais-Suomen maakuntaan, koska Ahvenanmaata koskevat tiedot voivat olla otoksen pienuuden vuoksi epäluotettavia. Tuloerojen ja terveyden yhteyttä tarkastellaan tutkielmassa myös seutukunnittain. Suomessa seutukunnat muodostavat NUTS 4-tason. Maakunnat ja seutukunnat luetellaan liitteessä 1.

Tutkimuksen aluepiirteitä kuvaavia muuttujia ovat alueen keskimääräinen tulotaso, alueiden sisäiset tuloerot ja köyhyysaste. Aluepiirteitä kuvaavien selittävien muuttujien kuvaukset esitetään liitteen 2 taulukossa 1. Aluetason muuttujien laskemisessa käytetään aineistona Tilastokeskuksen vuosien 2003 ja 2004 tulonjakotilastoja. Tutkielmassa maakunnan keskimääräinen tulo ja tuloeroja kuvaavat eriarvoisuusmitat lasketaan sekä kotitalouden jäsentä kohti (per capita) lasketun että ns. OECD-skaalan ja modifioidun OECD-skaalan avulla muunnetun kotitalouden käytettävissä olevan tulon perusteella. Skaalatut kotitalouden jäsenen käytettävissä olevat tulot painotetaan kotitalouden jäsenmäärällä, jotta tulojakauma muodostuisi yksilöistä kotitalouksien asemasta. Tiedot korotetaan perusjoukon tasolle kotitalouskohtaisten korotuskertoimien avulla. Tulomuuttujat muunnetaan alueellisia eriarvoisuus- ja köyhyysmittoja laskettaessa reaalisiksi. Perusvuotena käytetään vuotta 2004.

Vuoden 2003 tulonjakotilastossa yhden kotitalouden käytettävissä olevat tulot olivat poikkeuksellisen suuret. Lisäksi kotitalous sai suuren korotuskertoimen arvon. Maakunta- ja seutukuntakohtaisia eriarvoisuusmittoja laskettaessa kyseinen kotitalous tuotti yleiseen kehitykseen nähden poikkeuksellisen suuria Gini-kertoimen arvoja, joten kotitalous poistettiin aineistosta.

Maakunnan (seutukunnan)  $j$  keskimääräinen tulo  $\mu_j$  lasketaan:

$$\mu_j = \frac{\sum_{i=1}^n h_i w_i \left(\frac{y_i}{h_i}\right)}{\sum_{i=1}^n h_i w_i}, \quad (19)$$

jossa  $h_i$  on kotitalouden  $i$  jäsenmäärä maakunnassa (seutukunnassa)  $j$ ,  $w_i$  kotitalouskohtainen korotuskerroin ja  $y_i$  kotitalouden  $i$  käytettävissä oleva tulo. Osamäärä  $y_i/h_i$  kuvaa kotitalouden jäsentä kohti laskettua käytettävissä olevaa tuloa. Laskettaessa kotitalouden jäsenen ekvivalentti käytettävissä oleva tulo, jakajana käytetään jäsenmäärän sijaan edellä esitettyjä OECD-skaalaa tai modifioitua OECD-skaalaa. Myös tällöin kotitalouden  $i$  jäsenen ekvivalentti käytettävissä oleva tulo painotetaan jäsenmäärällä  $h_i$ . Hagforsin (1993, 415) mukaan kotitaloutta tulee painottaa jäsenmäärällä, vaikka tulokäsitteenä käytetään kotitalouden jäsenen ekvivalenttia tuloa, sillä niin sanottua ekvivalentti aikuista ei ole. Maakuntien ja seutukuntien keskimääräisen tulon kuvauksellinen analyysi esitetään liitteen 2 taulukoissa 2 ja 3.

Alueiden sisäiset tuloerot lasketaan luvussa 2.1.2 esitettyjen eriarvoisuusmittojen perusteella. Gini-kertoimen laskeminen perustuu yhtälöön 3, jossa  $\mu$  on maakunnan (seutukunnan)  $j$  keskimääräinen tulo  $\mu_j$  ja  $y$  joko yksilön ekvivalentti tai per capita käytettävissä oleva tulo. Gini-kertoimen laskemista varten kotitaloudet järjestetään maakunnittain (seutukunnittain) käytettävissä olevien tulojen perusteella nousevaan järjestykseen siten, että tuloja painotetaan sekä kotitalouden jäsenmäärällä että kotitalouskohtaisella korotuskertoimella. Atkinsonin indeksi lasketaan maakunnittain parametrin  $\varepsilon$  arvolla 1. Yleinen entropiamittari lasketaan puolestaan maakunnittain parametrin  $\theta$  arvolla 1, jolloin saadaan yleisen entropiamittarin erityistapaus Theilin indeksi. Theilin indeksin kaavassa 10 muuttuja  $\bar{y}$  on maakunnan  $j$  keskimääräinen tulotaso  $\mu_j$  (ks. kaava 19). Köyhyysmittoina käytetään maakunnittain laskettuja päälukumittaa (kaava 12) ja köyhyyskuilua (kaava 13). Köyhyysmitat korotetaan perusjoukon tasolla siten, että kaavoissa 12 ja 13  $(1/n)$  korvataan  $f_i = h_i w_i / N$ , jossa  $N = \sum_{i=1}^n h_i w_i$ . Alueellisten eriarvoisuus- ja köyhyysmittojen keskiarvo, keskihajonta, minimi ja maksimi vuosien 2003 ja 2004 tutkimusaineistossa esitetään liitteen 2 taulukoissa 2 ja 3.



### 5.3.3 Yksilön ja kotitalouden piirteitä kuvaavat selittävät muuttujat

Tutkielman yksilön ja kotitalouden piirteitä kuvaaviksi selittäviksi muuttujiksi on valittu pitkälti samoja selittäjiä kuin useissa aikaisemmin julkaistuissa tuloerojen ja terveyden välistä yhteyttä selvittäneissä tutkimuksissa käytetyt kontrollimuuttujat (ks. esim. Blomgren ym. 2004, Martikainen ym. 2004, Dahl ym. 2006, Li & Zhu 2006). Kyseessä olevat muuttujat ovat myös pitkälti samoja kuin luvussa 2.2.1 esitetyt terveyteen vaikuttavat tekijät. Muuttujien kuvaukset esitetään liitteen 2 taulukossa 1. Edellä mainitun liitteen taulukoissa 2 ja 3 raportoidaan puolestaan kyseessä olevien selittäjien keskiarvot, keskihajonnat, minimi ja maksimit.

Alla olevassa taulukossa 4 on kuvattu selittävien muuttujien keskiarvot ja keskihajonnat vuosien 2003 ja 2004 otoksessa niiden henkilöiden osalta, joilla oli pitkäaikaissairauksia ja jotka olivat terveitä. Taulukossa 5 esitetään puolestaan vastaavat luvut huonon tai melko huonon sekä hyvän yleisen terveydentilan omaaville henkilöille. Taulukoiden avulla pyritään kuvaamaan terveydentilaltaan erilaisten ryhmien välillä vallitsevia eroja.

Taulukosta 4 voidaan havaita, että pitkäaikaissairaat ovat keskimäärin vanhempia ja asuvat pienemmissä perheissä kuin henkilöt, jotka eivät ole sairaita. Pitkäaikaissairaiden ryhmässä on myös suhteellisesti enemmän eläkeläisiä, pitkäaikaistyöttömiä ja perusasteen koulutuksen suorittaneita kuin terveiden joukossa, jossa esimerkiksi korkea-asteen koulutuksen käyneitä ja ylempiä toimihenkilöitä on suhteellisesti enemmän. Hieman yllättäen pitkäaikaissairaiden käytettävissä olevat tulot ovat kuitenkin keskimäärin suuremmat kuin terveillä. Ryhmien ero on tilastollisesti merkitsevä ainoastaan vuonna 2004. Toisaalta terveiden ryhmässä maakunnan tulojen keskiarvo on keskimäärin suurempi molempina vuosina ja köyhyysaste keskimäärin hieman pienempi vuonna 2004 kuin pitkäaikaissairaila. Vuoden 2003 otoksessa köyhyysasteet eivät eroa tilastollisesti merkitsevästi ryhmien välillä. Maakunnan tuloerot ovat terveiden ryhmässä hieman korkeammat kuin pitkäaikaissairaiden ryhmässä kaikilla kolmella eriarvoisuusmitalla mitattuina. Seutukunnittain tuloerot ovat tilastollisesti merkitsevästi korkeammat terveiden ryhmässä ainoastaan vuoden 2004 aineistossa.

Taulukko 4. Selittävien muuttujien keskiarvo ja keskihajonta ( $\sigma$ ) otoksen henkilöillä, joilla ei ollut (0) ja joilla oli pitkäaikaissairauksia (1) vuosina 2003 ja 2004

	2003					2004				
	Pitkäaikais-sairastavuus = 0		Pitkäaikais-sairastavuus = 1		t-arvo	Pitkäaikais-sairastavuus = 0		Pitkäaikais-sairastavuus = 1		t-arvo
Havaintojen lukumäärä	6575		4254			6405		4506		
	Keskiarvo	σ	Keskiarvo	σ		Keskiarvo	σ	Keskiarvo	σ	
MUUTTUJAT <sup>a</sup>										
Selittävät muuttujat:										
Aluepiirteitä kuvaavat muuttujat <sup>b</sup>										
Keskimääräinen tulo1 <sup>c</sup>	0,144	0,016	0,142	0,016	6,35***	0,152	0,018	0,150	0,017	5,85***
Keskimääräinen tulo2 <sup>d</sup>	0,143	0,019	0,142	0,019	2,67***	0,152	0,021	0,150	0,021	4,90***
Gini-kerroin1 <sup>e</sup>	0,273	0,022	0,271	0,022	4,62***	0,279	0,027	0,277	0,027	3,81***
Gini-kerroin2 <sup>d</sup>	0,267	0,036	0,266	0,036	1,41	0,275	0,037	0,272	0,038	4,12***
Atkinsonin indeksi <sup>c</sup>	0,120	0,021	0,119	0,021	2,42**	0,127	0,026	0,125	0,026	3,96***
Theilinin indeksi <sup>c</sup>	0,151	0,053	0,148	0,053	2,88***	0,172	0,065	0,168	0,064	3,19***
Päälukumitta <sup>e</sup>	0,137	0,043	0,138	0,043	-1,18	0,128	0,047	0,130	0,047	-2,19**
Köyhyyskuilu <sup>c</sup>	0,030	0,011	0,030	0,011	0,00	0,027	0,012	0,028	0,012	-4,29***
Yksilöpiirteitä kuvaavat muuttujat										
Tulot <sup>b</sup>	0,166	0,198	0,168	0,151	-0,56	0,167	0,143	0,175	0,283	-1,94*
(Tulot <sup>2</sup> ) <sup>b</sup>	0,067	1,611	0,051	0,664	0,61	0,048	0,373	0,111	3,392	-1,47
QR <sup>e</sup>	1,769	0,882	1,839	0,816	-4,15***	1,768	0,916	1,830	0,844	-3,60***
Ikä	42,079	15,040	54,770	15,923	-41,90***	40,950	15,353	54,688	16,673	-44,41***
Ikä <sup>2</sup>	1996,85	1350,78	3253,20	1707,31	-42,54***	1912,54	1355,02	3268,66	1775,54	-45,21***
Sukupuoli	0,544	0,498	0,489	0,500	5,60***	0,508	0,500	0,475	0,499	3,40***
Siviilisäät	0,526	0,499	0,563	0,496	-3,78***	0,493	0,500	0,566	0,496	-7,53***
Koulutus:										
Perusaste	0,263	0,440	0,406	0,491	-15,77***	0,252	0,434	0,404	0,491	-17,05***
Keskiaste	0,420	0,494	0,361	0,480	6,14***	0,442	0,500	0,365	0,481	8,05***
Alin korkea-aste	0,142	0,349	0,122	0,327	2,98***	0,133	0,340	0,114	0,318	2,95***
Korkea-aste	0,174	0,379	0,111	0,314	9,02***	0,173	0,378	0,117	0,321	8,10***
Sosioekonominen asema:										
Ylempi toimihenkilö	0,157	0,363	0,091	0,287	10,01***	0,152	0,359	0,085	0,280	10,48***
Alempi toimihenkilö	0,167	0,373	0,105	0,307	9,04***	0,177	0,382	0,111	0,314	9,55***
Työntekijä	0,158	0,365	0,102	0,303	8,32***	0,159	0,365	0,096	0,295	9,59***
Yrittäjä	0,137	0,343	0,087	0,282	7,93***	0,120	0,325	0,075	0,263	7,69***
Maatalous	0,071	0,258	0,064	0,244	1,41	0,057	0,232	0,052	0,223	1,13
Opiskelija	0,108	0,311	0,040	0,195	12,73***	0,129	0,335	0,048	0,214	14,31***
Eläkeläinen	0,102	0,303	0,423	0,494	-41,90***	0,092	0,289	0,435	0,496	-45,45***
Pitkäaikaistyötön	0,057	0,232	0,064	0,244	-1,50	0,064	0,245	0,068	0,252	-0,83
Muu	0,043	0,204	0,025	0,155	4,91***	0,050	0,218	0,028	0,166	5,71***
Kotitalouden piirteitä kuvaavat muuttujat										
Kotitalouden koko	2,798	1,485	2,239	1,195	-21,58***	2,790	1,486	2,282	1,252	18,74***

<sup>a</sup> Muuttujien kuvaus liitteen 2 taulukossa 1

<sup>b</sup> Ekvivalenssisikaalana per capita

<sup>c</sup> Alueluokituksena maakunta

<sup>d</sup> Alueluokituksena seutukunta

<sup>e</sup> Q = Gini-kerroin1

$\sigma$  = Keskihajonta

Otoskeskiarvojen ero tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.

Taulukosta 5 voidaan havaita, että myös huonon tai melko huonon yleisen terveydentilan omaavat henkilöt ovat keskimäärin vanhempia ja asuvat pienemmissä perheissä kuin henkilöt, joiden terveydentila on hyvä. Huonon tai melko huonon terveydentilan omaavien ryhmässä on myös suhteellisesti enemmän eläkeläisiä, pitkäaikaistyöttömiä ja perusasteen koulutuksen suorittaneita ja toisaalta suhteellisesti vähemmän korkea-asteen koulutuksen käyneitä ja ylempiä toimihenkilöitä kuin niillä henkilöillä, joilla yleinen terveydentila on hyvä. Vuoden



Taulukko 5. Selittävien muuttujien keskiarvo ja keskihajonta ( $\sigma$ ) otoksen henkilöillä, joilla oli huono tai melko huono (0) sekä hyvä (3) yleinen terveydentila vuosina 2003 ja 2004

	2003					2004				
	Yleinen terveydentila = 3		Yleinen terveydentila = 0		t-arvo	Yleinen terveydentila = 3		Yleinen terveydentila = 0		t-arvo
Havaintojen lukumäärä	4846		957			4839		996		
	Keskiarvo	$\sigma$	Keskiarvo	$\sigma$		Keskiarvo	$\sigma$	Keskiarvo	$\sigma$	
<b>MUUTTUJAT<sup>a</sup></b>										
<i>Selittävät muuttujat:</i>										
<i>Aluepiirteitä kuvaavat muuttujat<sup>b</sup></i>										
Keskimääräinen tulo <sup>c</sup>	0,145	0,016	0,141	0,015	7,14***	0,153	0,018	0,150	0,017	4,83***
Keskimääräinen tulo <sup>d</sup>	0,145	0,019	0,140	0,019	7,44***	0,153	0,021	0,149	0,021	5,47***
Gini-kerroin <sup>e</sup>	0,274	0,022	0,271	0,022	3,86***	0,280	0,027	0,276	0,027	4,26***
Gini-kerroin <sup>d</sup>	0,269	0,035	0,265	0,037	3,20***	0,277	0,037	0,270	0,037	5,44***
Atkinsonin indeksi <sup>e</sup>	0,121	0,021	0,119	0,021	2,69***	0,128	0,026	0,124	0,026	4,42***
Theilin indeksi <sup>e</sup>	0,154	0,052	0,148	0,054	3,24***	0,175	0,065	0,166	0,063	4,00***
Päälukumitta <sup>e</sup>	0,135	0,043	0,139	0,043	-2,63***	0,127	0,047	0,130	0,046	-1,84*
Köyhyyskuilu <sup>e</sup>	0,029	0,011	0,031	0,012	-5,06***	0,027	0,012	0,028	0,012	-2,40***
<i>Yksilöpiirteitä kuvaavat muuttujat</i>										
Tulot <sup>b</sup>	0,169	0,225	0,159	0,219	1,26	0,174	0,290	0,154	0,076	2,16***
(Tulot <sup>2</sup> ) <sup>b</sup>	0,079	1,884	0,073	1,299	0,09	0,114	3,287	0,029	0,041	0,82
QR <sup>e</sup>	1,773	0,892	0,543	0,837	39,37***	1,771	0,928	0,506	0,777	40,22***
Ikä	39,672	14,519	60,188	14,987	-39,73***	38,665	14,973	60,462	15,224	-41,72***
Ikä <sup>2</sup>	1784,61	1254,70	3846,99	1776,88	-43,04***	1719,14	1282,79	3887,18	1798,07	-45,01***
Sukupuoli	0,518	0,500	0,519	0,500	-0,06	0,490	0,500	0,494	0,500	-0,23
Siviilisäätö	0,494	0,500	0,515	0,500	-1,19	0,472	0,499	0,529	0,499	-3,28***
<i>Koulutus:</i>										
Perusaste	0,230	0,421	0,516	0,500	-18,59***	0,228	0,419	0,515	0,500	-19,01***
Keskiaste	0,417	0,493	0,320	0,467	5,61***	0,441	0,497	0,341	0,474	5,83***
Alin korkea-aste	0,150	0,357	0,099	0,299	4,14***	0,138	0,348	0,077	0,267	5,22***
Korkea-aste	0,203	0,402	0,065	0,246	10,25***	0,193	0,395	0,066	0,249	9,76***
<i>Sosioekonominen asema:</i>										
Ylempi toimihenkilö	0,183	0,386	0,032	0,177	11,86***	0,169	0,375	0,035	0,184	11,01***
Alempi toimihenkilö	0,181	0,385	0,065	0,246	9,00***	0,189	0,392	0,053	0,225	10,60***
Työntekijä	0,149	0,356	0,063	0,243	7,15***	0,147	0,354	0,055	0,229	7,87***
Yrittäjä	0,121	0,326	0,065	0,246	5,04***	0,106	0,308	0,056	0,230	4,85***
Maatalousyrittäjä	0,053	0,225	0,057	0,233	-0,50	0,045	0,207	0,045	0,208	0,00
Opiskelija	0,132	0,338	0,008	0,091	11,27***	0,158	0,364	0,019	0,137	11,88***
Eläkeläinen	0,078	0,268	0,608	0,488	-47,56***	0,074	0,262	0,614	0,487	-49,73***
Pitkäaikaistyötön	0,054	0,227	0,070	0,255	-1,95*	0,059	0,235	0,088	0,284	-3,42***
Muu	0,049	0,215	0,031	0,174	2,44**	0,053	0,223	0,033	0,179	2,66***
<i>Kotitalouden piirteitä kuvaavat muuttujat</i>										
Kotitalouden koko	2,833	1,520	2,070	1,159	14,71***	2,826	1,497	2,052	1,115	15,46***

<sup>a</sup> Muuttujien kuvaus liitteen 2 taulukossa 1

<sup>b</sup> Ekvivalenssisikaalana per capita

<sup>c</sup> Alueluokituksena maakunta

<sup>d</sup> Alueluokituksena seutukunta

<sup>e</sup> Q = Gini-kerroin

$\sigma$  = Keskihajonta

Otoskeskiarvojen ero tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.

2004 otoksessa käytettävissä olevat tulot ovat odotetusti keskimäärin suuremmat hyvän yleisen terveydentilan omaavilla henkilöillä kuin niillä, joiden yleinen terveydentila oli huono tai melko huono. Myös vuoden 2003 aineistossa tulot ovat suuremmat hyvän terveydentilan omaavilla, mutta otoskeskiarvojen ero ryhmien välillä ei ole tällöin tilastollisesti merkitsevä. Odotusten mukaisesti maakunnan keskiarvotulo on keskimäärin suurempi ja köyhyysaste pienempi hyvän terveydentilan omaavien ryhmässä. Maakunnan tuloerot ovat kuitenkin ennako-odotusten vastaisesti hieman korkeammat edellä mainitussa ryhmässä kaikilla

kolmella eriarvoisuusmitalla mitattuina. Vastaava havainto voidaan tehdä myös seutukuntien tuloerojen osalta. Alueellisten eriarvoisuus- ja köyhyysmittojen osalta erot ryhmien välillä eivät ole kuitenkaan merkittävän suuret vaikkakin tilastollisesti merkitsevät.

#### 5.4 Tutkimusmenetelmät

Kuten edellä havaittiin, tutkielman selitettävät muuttujat eivät ole jatkuvia, vaan ne saavat tiettyjä diskreettejä arvoja. Lineaarisia regressiomalleja ei voida soveltaa estimoinnissa, mikäli riippuva muuttuja ei voi saada mitä tahansa arvoja (Greene 2003, 663). Käsillä olevassa tutkielmassa joudutaan näin ollen käyttämään epälineaarisia malleja.

Tarkastellaan selitettävien muuttujien lähtökohtana olevia valintajoukkoja. Merkitään valintajoukkoa  $I = 0, \dots, J$ . Vastemuuttuja on binäärinen, kun valintaluokkia on kaksi (Greene 2003, 665). Tällöin  $J = 1$ . Vastemuuttuja on puolestaan polytonominen, kun valintaluokkia on enemmän kuin kaksi (Maddala 1983, 34). Tällöin  $J > 1$ . Tässä tutkielmassa mallinnetaan sekä tilannetta, jolloin valintaluokkia on kaksi (pitkäaikaissairastavuus) että tilannetta, jolloin valintaluokkia on neljä (yleinen terveydentila). Ensimmäisessä tapauksessa valintajoukko on  $I = 0, 1$  ja jälkimmäisessä tapauksessa  $I = 0, \dots, 3$  (vrt. liite 2, taulukko 1).

Malleja, joissa riippuva muuttuja saa diskreettejä arvoja, kutsutaan tässä diskreetin valinnan malleiksi (vrt. Greene 2003, 663). Tutkielman malleissa riippuvat muuttujat kuvaavat todennäköisyyttä, että henkilö kuuluu terveydentilaltaan luokkaan  $j$   $J$ :stä toisensa poissulkevasta vaihtoehdosta. Tarkastelu aloitetaan kaksiarvoisen diskreetin valinnan mallista ja sen erityistapauksesta probit-mallista. Tässä mallissa vastemuuttuja on binäärinen ja kuvaa käsillä olevassa tutkielmassa henkilön pitkäaikaissairastavuuden todennäköisyyttä. Tämän jälkeen tarkastelu laajennetaan diskreetin valinnan malliin, jossa lähtökohtana on usean arvon ordinaalinen valintajoukko. Tällöin mallinnetaan henkilön yleistä terveydentilaa. Valintajoukko on ordinaalinen, koska yleistä terveydentilaa kuvaavat vaihtoehdot voidaan asettaa luonnolliseen järjestykseen (yleinen terveydentila on huono, melko huono, keskitasoinen jne.). Tutkielmassa yleisen terveydentilan mallintamiseen sovelletaan probit-mallin yleistystä, järjestettyä probit -mallia. Seuraavassa käydään läpi tutkielman ekonometrisessa analyysissä käytettävät mallit, estimointimenetelmät ja estimoitavien parametrien tulkinta.



### 5.4.1 Probit-malli

Kaksiarvoisen diskreetin valinnan todennäköisyysmallin tarkastelu voidaan aloittaa kuvaamalla tapahtuman todennäköisyyttä ja sen komplementtia. Tässä tapahtuman todennäköisyydellä tarkoitetaan henkilön  $i=1, \dots, n$  pitkäaikaissairastavuuden todennäköisyyttä:

$$P(H_i = 1 | X_i) = F(X_i, \beta) \quad (20)$$

$$P(H_i = 0 | X_i) = 1 - F(X_i, \beta), \quad (21)$$

jossa  $H_i$  saa arvon 1, jos henkilöllä  $i$  on pitkäaikaissairauksia ja 0, jos ei.  $X_i$  on selittävien muuttujien vektori ja parametrivektori  $\beta$  kuvaa vektorissa  $X_i$  tapahtuvien muutosten vaikutusta pitkäaikaissairastavuuden todennäköisyyteen  $P$  (vrt. Greene 2003, 665).

Funktioon  $F(X_i, \beta)$  voidaan soveltaa eri funktiomuotoja riippuen tehdyistä lähtöoletuksista. Funktiona voidaan käyttää esimerkiksi lineaarista todennäköisyysmallia  $F(X_i, \beta) = \beta X_i$  (Greene 2003, 665). Mallin käyttöön liittyy kuitenkin ongelmia: virhetermi on heteroskedastinen<sup>22</sup> eivätkä mallin ennusteet rajoitu todennäköisyyden  $P$  vaatimien arvojen 0 ja 1 välillä. Joudutaan siis käyttämään sellaista funktion  $F(X_i, \beta)$  muotoa, jossa mallin ennusteet ovat välillä  $[0,1]$ . (mt. 665–666) Tällöin voidaan käyttää probit- ja logit-malleja.

Seuraavassa käsitellään henkilön  $i$  pitkäaikaissairastavuuden todennäköisyyttä selittävää probit-mallia. Yhtälössä 22  $H_i^*$  on havaitsematon vastemuuttuja, joka on lineaarisesti riippuvainen selittävien muuttujien vektorista  $X_i$ . Tässä tutkielmassa selittävät muuttujat kuvaavat henkilön  $i$  piirteiden lisäksi kyseessä olevan henkilön asuinalueen ja kotitalouden piirteitä. Muuttujat esiteltiin edellä luvuissa 5.3.2 ja 5.3.3. Yhtälössä 22  $\beta'$  on tuntematon parametrivektori, ja virhetermi  $\varepsilon_i$  kuvaa selittämätöntä satunnaisvaihtelua. Muuttuja  $H_i$

<sup>22</sup> Merkitään lineaarista todennäköisyysmallia  $F(X_i, \beta) = \beta'X_i$ . Koska  $E(H_i|X_i) = F(X_i, \beta)$ , on  $H_i = E(H_i|X_i) + \varepsilon_i = \beta'X_i + \varepsilon_i$ .  $H_i$  voi saada tässä arvoja 0 tai 1, joten virhetermi  $\varepsilon_i$  on näin ollen joko  $-\beta'X_i$  tai  $1-\beta'X_i$  todennäköisyyksillä  $F$  tai  $1-F$ . Virhetermin varianssi on tällöin  $\sigma^2 = [(-\beta'X_i)^2 F] + [(1-\beta'X_i)^2 (1-F)] = \beta'X_i(1-\beta'X_i)$ . Virhetermi on heteroskedastinen, koska sen varianssi ei ole vakio kaikilla muuttujan  $X_i$  arvoilla. (Greene 2003, 665)

yhtälöissä 23 ja 24 on puolestaan havaitsemattoman muuttujan  $H_i^*$  indikaattorimuuttuja, joka saa arvon 1, kun henkilöllä  $i=1, \dots, n$  on pitkäaikaissairauksia, ja 0, jos ei:

$$H_i^* = \beta'X_i + \varepsilon_i, \quad (22)$$

$$H_i = 1, \text{ jos } H_i^* > 0, \quad (23)$$

$$H_i = 0, \text{ jos } H_i^* \leq 0. \quad (24)$$

Henkilön  $i$  pitkäaikaissairastavuuden todennäköisyys voidaan esittää näin ollen:

$$\begin{aligned} P(H_i = 1 | X_i) &= P(H_i^* > 0 | X_i) \\ &= P(\beta'X_i + \varepsilon_i > 0 | X_i) \\ &= P(\varepsilon_i > -\beta'X_i | X_i) \\ &= 1 - F(-\beta'X_i) \\ &= F(\beta'X_i), \end{aligned} \quad (25)$$

jossa funktio  $F$  on virhetermin  $\varepsilon_i$  kertymäfunktio. (vrt. esim. Maddala 1983, 22–23)

Probit-mallissa virhetermit  $\varepsilon_i$  oletetaan normaalisti jakautuneiksi,  $\varepsilon_i \sim N(0,1)$ . Tällöin virhetermin  $\varepsilon_i$  kertymäfunktio  $F$  korvataan standardoidun normaalijakauman kertymäfunktioilla  $\Phi(\cdot)$ . (Maddala 1983, 22–23) Henkilön  $i$  pitkäaikaissairastavuuden todennäköisyys on tällöin:

$$P(H_i = 1 | X_i) = \int_{-\infty}^{\beta'X_i} \phi(t) dt = \Phi(\beta'X_i), \quad (26)$$

jossa  $\phi(t)$  on normaalijakauman tiheysfunktio. Logit-mallissa virhetermien oletetaan olevan logistisesti jakautuneet ja virhetermin  $\varepsilon_i$  kertymäfunktio  $F$  korvataan logistisen jakauman kertymäfunktioilla. (Greene 2003, 666–667) Käytännössä logistinen ja normaalijakauma tuottavat lähes samanlaisia tuloksia (Greene 2003, 737), eikä mallin valinnalla ole useinkaan



tulosten kannalta merkitystä (mt. 667). Probit-mallia on käytetty monissa tuloerojen ja terveyden välistä yhteyttä tarkastelleissa tutkimuksissa (ks. Mellor & Milyo 2002, Gertdham & Johannesson 2004, Hildebrandt & Van Kerm 2005, Li & Zhu 2006).

Probit-mallin epälineaarisuuden takia parametriestimaatteja ei voida tulkita marginaalivaikutuksiksi, kuten lineaarisessa regressiomallissa (Greene 2003, 667–668). Oletetaan, että  $X_{ik}$  ja  $\beta_k$  muodostavat selittävien muuttujien vektorin  $X_i$  ja parametrinvektorin  $\beta$  k:nneen elementin. Kertoimesta  $\beta_k$  ei voida tehdä tässä tapauksessa päätelmiä siitä, kuinka paljon selittäjän  $X_{ik}$  arvon muutokset vaikuttavat henkilön  $i$  pitkäaikaissairastavuuden todennäköisyyteen. Tilannetta voidaan havainnollistaa laskemalla todennäköisyyksien derivaatat  $X_{ik}$  suhteen. Merkitään todennäköisyysfunktion derivaattaa jatkuvan muuttuja  $X_{ik}$  suhteen yleisesti:

$$\frac{E[H_i | X_i]}{\partial X_{ik}} = \frac{\partial F(\beta'X_i)}{\partial(\beta'X_i)} \times \frac{\partial(\beta'X_i)}{\partial X_{ik}} = f(\beta'X_i)\beta_k, \quad (27)$$

jossa  $f(\cdot)$  on kertymäfunktioita  $F(\cdot)$  vastaava tiheysfunktio (mt. 667). Probit-mallin tapauksessa  $E[H_i | X_i] = 0 \times [1 - \Phi(\beta'X_i)] + 1 \times [\Phi(\beta'X_i)] = \Phi(\beta'X_i)$ , joten todennäköisyysfunktion osittaisderivaataksi  $X_{ik}$  suhteen saadaan:

$$\frac{E[H_i | X_i]}{\partial X_{ik}} = \frac{\partial \Phi(\beta'X_i)}{\partial(\beta'X_i)} \times \frac{\partial(\beta'X_i)}{\partial X_{ik}} = \phi(\beta'X_i)\beta_k, \quad (28)$$

jossa  $\phi(\cdot)$  on normaalijakauman tiheysfunktio (mt. 667–668).

Kaavasta 28 nähdään, että  $\beta_k$  ei yksistään selitä  $X_{ik}$ :n vaikutusta henkilön  $i$  pitkäaikaissairastavuuden todennäköisyyteen, vaan siihen vaikuttavat myös termin  $\phi(\beta'X_i)$  välityksellä selittävien muuttujien vektori  $X_i$  ja parametrivektori  $\beta'$ . Koska normaalijakauman tiheysfunktio kuvaa pistetodennäköisyyttä eikä todennäköisyys voi saada negatiivisia arvoja, voidaan kertoimen  $\beta_k$  etumerkistä kuitenkin päätellä, mihin suuntaan todennäköisyys muuttuu selittäjän  $X_{ik}$  arvossa tapahtuvien muutosten seurauksena. Mikäli

kertoimen  $\beta_k$  etumerkki on positiivinen, tapahtuman todennäköisyys kasvaa selittävän muuttujan  $X_{ik}$  arvon kasvaessa. Kertoimen  $\beta_k$  ollessa negatiivinen selittäjän  $X_{ik}$  arvon kasvaminen pienentää todennäköisyyttä. Selittävien muuttujien kertoimia vertaamalla saadaan selville myös, millä muuttujalla on suurin vaikutus pitkäaikaissairastavuuden todennäköisyyteen. Mitä suurempi kertoimen arvo on, sitä enemmän todennäköisyys kasvaa selittäjän arvon noustessa.

Kertoimesta  $\beta_k$  ei voida päätellä, kuinka paljon tapahtuman todennäköisyys muuttuu selittäjän  $X_{ik}$  arvossa tapahtuvien muutosten seurauksena. Mikäli halutaan tietää, kuinka paljon henkilön  $i$  pitkäaikaissairastavuuden todennäköisyys muuttuu selittäjän  $X_{ik}$  arvon kasvaessa yhdellä yksiköllä tai dummy-muuttujan kasvaessa nolasta arvoon yksi, joudutaan laskemaan marginaalivaikutus edellä esitetyn kaavan 28 mukaisesti. Koska todennäköisyysfunktion osittaisderivaatta  $X_{ik}$ :n suhteen vaihtelee tiheysfunktion  $\phi(\beta'X_i)$   $\beta'X_i$ :n eri arvoilla, lasketaan marginaalivaikutukset yleensä selittävien muuttujien keskiarvopisteessä (vrt. Greene 2003, 668). Tällöin selittävien muuttujien vektori  $X_i$  korvataan keskiarvovektorilla  $\bar{X}$ .

#### 5.4.2 Järjestetty probit -malli

Tutkimusmenetelmistä on käsitelty tähän mennessä tuloerojen ja terveyden välisen yhteyden mallintamista, kun selitettävänä muuttujana on henkilön pitkäaikaissairastavuus. Tutkielman toinen selitettävä muuttuja on yleinen terveydentila. Tällöin tuloerojen ja terveyden välisen yhteyden mallintamisessa ei voida käyttää tavallista probit-mallia, koska yleinen terveydentila -muuttujan lähtökohtana on usean arvon ordinaalinen valintajoukko. Joudutaan siis soveltamaan moniluokkaiselle kategoriselle muuttujalle tarkoitettua diskreetin valinnan mallia, jossa lähtökohtana on ordinaalinen valintajoukko. Käsillä olevassa tutkielmassa yleisen terveydentilan mallintamiseen sovelletaan probit-mallin yleistystä järjestettyä probit -mallia (engl. ordered probit model).

Järjestetyn probit -mallin käsittely voidaan aloittaa yhtälöstä 29, jossa  $H_i^*$  on lineaarisesti selittävien muuttujien vektorista  $X_i$  riippuva havaitsematon vastemuuttuja.  $H_i^*$  kuvaa nyt



henkilön  $i$  havaitsematonta yleistä terveydentilaa. Selittävät muuttujat  $X_i$  ovat samoja kuin edellä. Yhtälössä 29  $\beta'$  on tuntematon parametrivektori, ja virhetermi  $\varepsilon_i$  kuvaa selittämätöntä satunnaisvaihtelua:

$$H_i^* = \beta'X_i + \varepsilon_i. \quad (29)$$

Muuttuja  $H_i$  yhtälössä 30 on puolestaan havaitsemattoman muuttujan  $H_i^*$  indikaattorimuuttuja, joka tässä tutkielmassa saa arvon 0, kun henkilön  $i=1, \dots, n$  yleinen terveydentila on melko huono tai huono, 1 jos keskitasoinen, jne. (vrt. liite 2, taulukko 1):

$$\begin{aligned} H_i &= 0 && \text{jos } H_i^* \leq 0, \\ H_i &= 1 && \text{jos } 0 < H_i^* \leq \mu_1, \\ H_i &= 2 && \text{jos } \mu_1 < H_i^* \leq \mu_2, \\ H_i &= 3 && \text{jos } \mu_2 > H_i^*, \end{aligned} \quad (30)$$

jossa  $\mu_1$  ja  $\mu_2$  ovat  $\beta$ :n kanssa estimoitavia tuntemattomia parametreja. (vrt. Greene 2003, 736)

Kuten edellä probit-mallin tapauksessa oletetaan seuraavassa, että virhetermit  $\varepsilon_i$  ovat normaalisti jakautuneet. Kunkin henkilön  $i$  yleistä terveydentilaa kuvaavan vaihtoehdon  $H_i = 0, 1, \dots, 3$  todennäköisyys voidaan määritellä tällöin:

$$\begin{aligned} P(H_i = 0 | X_i) &= P(H_i^* \leq 0 | X_i) = P(\beta'X_i + \varepsilon_i \leq 0 | X_i) = \Phi(-\beta'X_i), \\ P(H_i = 1 | X_i) &= P(0 < H_i^* \leq \mu_1 | X_i) = \Phi(\mu_1 - \beta'X_i) - \Phi(-\beta'X_i), \\ P(H_i = 2 | X_i) &= P(\mu_1 < H_i^* \leq \mu_2 | X_i) = \Phi(\mu_2 - \beta'X_i) - \Phi(\mu_1 - \beta'X_i), \\ P(H_i = 3 | X_i) &= P(H_i^* > \mu_2 | X_i) = 1 - \Phi(\mu_2 - \beta'X_i), \end{aligned} \quad (31)$$

jossa  $0 < \mu_1 < \mu_2$ . Mikäli vaihtoehtoja olisi kaksi, malli palautuisi perinteiseksi probit-malliksi. (vrt. Greene 2003, 737)

Probit-mallin tavoin myöskään järjestetyn probit -mallin parametriestimaatteja ei voida tulkita marginaalivaikutuksiksi mallin epälineaarisuuden takia. Merkitään selittävien muuttujien vektorin  $X_i$  ja parametrinvektorin  $\beta$  k:nnessa elementtiä  $X_{ik}$  ja  $\beta_k$ . Marginaalivaikutukset saadaan derivoimalla kaava 31 selittävän muuttujan  $X_{ik}$  suhteen (vrt. Greene 2003, 738):

$$\begin{aligned}\frac{\partial P(H_i = 0 | X_i)}{\partial X_{ik}} &= -\phi(-\beta'X_i)\beta_k, \\ \frac{\partial P(H_i = 1 | X_i)}{\partial X_{ik}} &= [\phi(-\beta'X_i) - \phi(\mu_1 - \beta'X_i)]\beta_k, \\ \frac{\partial P(H_i = 2 | X_i)}{\partial X_{ik}} &= [\phi(\mu_1 - \beta'X_i) - \phi(\mu_2 - \beta'X_i)]\beta_k, \\ \frac{\partial P(H_i = 3 | X_i)}{\partial X_{ik}} &= \phi(\mu_2 - \beta'X_i)\beta_k.\end{aligned}\tag{32}$$

Myös järjestetyn probit -mallin kohdalla marginaalivaikutukset riippuvat selittävien muuttujien  $X_i$  arvoista. Käsillä olevassa tutkielmassa marginaalivaikutukset lasketaan selittävien muuttujien  $X_i$  keskiarvopisteessä. Tällöin selittävien muuttujien vektori  $X_i$  korvataan kaavassa 32 selittävien muuttujien keskiarvovektorilla  $\bar{X}$ . Dummy-muuttujien tapauksessa marginaalimuutos kuvaa todennäköisyyden muutosta selittävän muuttujan arvon kasvaessa nolasta arvoon yksi muiden selittäjien ollessa keskiarvossaan. Järjestetyn probit-mallin tapauksessa marginaalivaikutusten tulkintaa vaikeuttaa se, että kertoimella  $\beta_k$  on yksiselitteinen tulkinta ainoastaan ylimmän ja alimman vastauskategorian (tässä: yleinen terveydentila on 0 tai 3) kohdalla, mikä voidaan havaita kaavasta 32 (vrt. Wooldridge 2002, 506). Näin ollen tutkielman tuloksia raportoidessa keskitytään järjestetyn probit -mallin tapauksessa marginaalivaikutusten tarkasteluun ylimpään ja alimpaan vastauskategoriaan.



## 6 TULOKSET

Luvussa esitetään probit- ja järjetetty probit –mallien estimointitulokset<sup>23</sup>. Ensimmäisessä alaluvussa käsitellään tuloerojen yhteyttä pitkäaikaissairastavuuteen. Toisessa alaluvussa tarkastelun kohteena on yleisen terveydentilan ja tuloerojen välisestä yhteydestä saadut estimointitulokset. Kunkin osion alussa käydään lyhyesti läpi tulosten esitystapa sekä tulosten tulkinnassa käytetyt periaatteet. Tämän jälkeen keskitytään estimointitulosten raportointiin ja tulkintaan.

### 6.1 Tuloerojen yhteys pitkäaikaissairastavuuteen

Taulukoissa 6 ja 7 esitetään probit-mallin estimointitulokset kolmelle tuloerojen ja pitkäaikaissairastavuuden välistä yhteyttä kuvaavalle yhtälölle. Mallissa 1 pitkäaikaissairastavuutta selitetään maakuntakohtaisella Gini-kertoimella. Malli 2 on muodostettu lisäämällä ensimmäiseen malliin maakunnan keskimääräinen tulo -muuttuja sekä yksilö- ja kotitalouden piirteitä kuvaavat kontrollimuuttujat. Malli 2 testaa tulojen eriarvoisuuden hypoteesin vahvaa versiota (tutkielman hypoteesi 1, ks. luku 3.2 yhtälö 17). Mallin 3 avulla selvitetään puolestaan, onko maakunnittain laskettujen tuloerojen yhteys yksilöiden pitkäaikaissairastavuuteen erilainen eri tulotasoilla olevien henkilöihin välillä. Malli on muodostettu lisäämällä muuttujat  $R$  ja  $QR$  malliin 2. Kolmannen mallin avulla testataan tulojen eriarvoisuuden hypoteesin heikkoa versiota (tutkielman hypoteesi 2, ks. luku 3.2 yhtälö 18).

Kuten edellä luvussa 5.4.1 havaittiin, probit-mallin positiivinen kerroinestimaatti merkitsee, että kyseessä olevan selittävän muuttujan arvon kasvaessa henkilön pitkäaikaissairastavuuden todennäköisyys nousee. Negatiivinen kerroinestimaatti on puolestaan merkki selittäjän pitkäaikaissairastavuuden todennäköisyyttä pienentävästä vaikutuksesta. Mikäli tulokset tukevat tulojen eriarvoisuuden hypoteesin vahvaa versiota, tuloeroja kuvaavan muuttujan kerroinestimaatti on positiivinen ja tilastollisesti merkitsevä. Vastaavasti  $QR$ -muuttujan kerroinestimaatin tulee olla negatiivinen ja tilastollisesti merkitsevä, mikäli tulokset tukevat tulojen eriarvoisuuden hypoteesin heikkoa versiota.

---

<sup>23</sup> Estimoinnit suoritettiin Stata/SE 9.2-ohjelmistolla.

Taulukko 6. Pitkäaikaissairastavuuden todennäköisyyden probit-mallin kerroinestimaatit vuonna 2003 (tuloerot laskettu maakunnittain)

Muuttuja	Malli 1			Malli 2			Malli 3		
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe
Gini-kerroin1	-2,365	1,288*	0,560	0,748	0,821	0,868	-0,628	1,804	1,803
Keskimääräinen tulo1				-2,026	1,266	1,205	-2,320	1,313*	1,250
R							-0,073	0,054	0,057
QR							0,233	0,192	0,206
Tulot				-0,169	0,090*	0,134	-0,039	0,104	0,156
Tulot <sup>2</sup>				0,011	0,014	0,018	-0,004	0,016	0,022
Ikä				0,051	0,004***	0,005	0,052	0,004***	0,005
Ikä <sup>2</sup>				0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli				-0,114	0,028***	0,027	-0,112	0,028***	0,027
Siviilisaäty				-0,076	0,020***	0,032	-0,072	0,019***	0,032
Keskiaste				-0,037	0,035	0,0330	-0,035	0,035	0,033
Alin korkea-aste				-0,097	0,043**	0,046	-0,092	0,043**	0,046
Korkea-aste				-0,255	0,037***	0,050	-0,250	0,036***	0,050
Ylempi toimihenkilö				0,056	0,065	0,059	0,057	0,066	0,059
Alempi toimihenkilö				-0,044	0,057	0,052	-0,041	0,057	0,052
Yrittäjä				-0,074	0,044*	0,053	-0,078	0,044*	0,053
Maatalousyrittäjä				0,126	0,054**	0,061	0,121	0,054**	0,061
Opiskelija				0,182	0,066***	0,072	0,175	0,066***	0,072
Eläkeläinen				0,702	0,051***	0,058	0,697	0,051***	0,058
Pitkäaikaistyötön				0,217	0,073***	0,063	0,199	0,077***	0,063
Muu				0,139	0,066**	0,081	0,128	0,070*	0,081
Kotitalouden koko				-0,065	0,014***	0,012	-0,070	0,014***	0,012
Vakio	0,371	0,339	0,153	-1,627	0,223***	0,221	-1,178	0,385***	0,459
Havaintojen lukumäärä	10829			10829			10829		
Pseudo R <sup>2</sup>	0,001			0,140			0,141		
Ennustettu oikein (%)	60,72			71,11			71,11		

Koulutus-muuttujassa vertailuryhmänä ovat perusasteen koulutuksen suorittaneet henkilöt ja sosioekonominen asema -muuttujassa työntekijät. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.

Taulukko 7. Pitkäaikaissairastavuuden todennäköisyyden probit-mallin kerroinestimaatit vuonna 2004 (tuloerot laskettu maakunnittain)

Muuttuja	Malli 1			Malli 2			Malli 3		
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe
Gini-kerroin1	-1,615	0,571***	0,450	0,026	0,786	0,664	0,695	1,222	1,287
Keskimääräinen tulo1				-0,060	1,227	1,025	-0,208	1,321	1,061
R							0,023	0,041	0,045
QR							-0,089	0,144	0,161
Tulot				-0,281	0,240	0,156	-0,242	0,182	0,185
Tulot <sup>2</sup>				0,072	0,046	0,046	0,065	0,032**	0,046
Ikä				0,045	0,005***	0,005	0,045	0,005***	0,005
Ikä <sup>2</sup>				0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli				-0,047	0,031	0,027	-0,046	0,031	0,027
Siviilisaäty				-0,023	0,027	0,033	-0,023	0,027	0,033
Keskiaste				-0,054	0,029*	0,033	-0,054	0,029*	0,033
Alin korkea-aste				-0,126	0,044***	0,047	-0,127	0,044***	0,047
Korkea-aste				-0,217	0,037***	0,051	-0,216	0,038***	0,051
Ylempi toimihenkilö				0,036	0,052	0,060	0,038	0,053	0,060
Alempi toimihenkilö				-0,011	0,048	0,051	-0,011	0,048	0,051
Yrittäjä				-0,077	0,043*	0,055	-0,078	0,042*	0,055
Maatalousyrittäjä				0,148	0,066**	0,064	0,147	0,066**	0,064
Opiskelija				0,177	0,080**	0,066	0,176	0,079**	0,067
Eläkeläinen				0,807	0,038***	0,059	0,806	0,037***	0,059
Pitkäaikaistyötön				0,270	0,045***	0,061	0,268	0,046***	0,061
Muu				0,139	0,066**	0,076	0,136	0,065**	0,076
Kotitalouden koko				-0,047	0,014***	0,011	-0,047	0,014***	0,012
Vakio	0,229	0,160	0,126	-1,620	0,272***	0,197	-1,776	0,283***	0,353
Havaintojen lukumäärä	10911			10911			10911		
Pseudo R <sup>2</sup>	0,001			0,152			0,152		
Ennustettu oikein (%)	58,70			71,41			71,41		

Koulutus-muuttujassa vertailuryhmänä ovat perusasteen koulutuksen suorittaneet henkilöt ja sosioekonominen asema -muuttujassa työntekijät. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.



Kerroinestimaattien perusteella ei voida kuitenkaan tehdä päätelmiä selittävän muuttujan vaikutuksen suuruudesta pitkäaikaissairastavuuden todennäköisyyteen. Tämän vuoksi taulukoissa 8 ja 9 esitetään vuosien 2003 ja 2004 tutkimusaineistosta lasketut marginaalivaikutukset. Kun taulukoissa 8 ja 9 esitetyt marginaalivaikutukset kerrotaan luvulla sata, voidaan luvut tulkita pitkäaikaissairastavuuden todennäköisyyden muutoksena prosenttiyksikössä, kun jatkuva selittäjä kasvaa yhdellä yksiköllä tai dummy-muuttujan nollasta arvoon yksi muiden selittäjien olleessa keskiarvossaan. Gini-kertoimen, alueen keskimääräisen tulon ja yksilöiden käytettävissä olevien tulojen ollessa tarkasteltavana muuttujana on tulosten tulkinnassa huomioitava muuttujan skaala. Koska Gini-kerroin saa arvoja nollasta yhteen, on tuloksia tulkittaessa mielekkäämpää tarkastella esimerkiksi kertoimen ensimmäisessä desimaalissa tapahtuvia muutoksia. Gini-kertoimen kasvaessa 0,10 yksiköllä, kerrotaan marginaalivaikutus luvun 100 sijaan luvulla 10. Koska alueen keskimääräinen tulo ja yksilöiden käytettävissä olevat tulot jaettiin  $10^5$  eurolla, kuvaa kyseessä olevien muuttujien marginaalivaikutusten sadalla kertominen pitkäaikaissairastavuuden todennäköisyyden muutosta prosenttiyksiköissä, kun tulot kasvavat  $10^5$  euroa. Tässäkin tapauksessa on mielekkäämpää tarkastella tulojen nousua esimerkiksi tuhannella eurolla. Tällöin keskimääräinen tulo- ja tulot-muuttujan marginaalivaikutus kerrotaan luvulla yksi, eli raportoidut marginaalivaikutukset voidaan tulkita prosenttiyksiköinä sellaisenaan.

Virhetermien laskemisessa otetaan huomioon, että maakunnan ja seutukunnan sisällä oletus havaintojen riippumattomuudesta ei välttämättä päde. Henkilöiden asuessa samalla alueella voivat esimerkiksi asuinalueen elinolosuhteet vaikuttaa terveyteen samansuuntaisesti. Tällöin oletus kahden virhetermin tulon odotusarvosta ei ole 0 ja keskivirheet voivat olla harhaisia. Tämän vuoksi selittäville muuttujille lasketaan keskivirheet, joissa havaintojen klusterointi on otettu huomioon. Jokaiselle selittäjälle ilmoitetaan taulukoissa myös korjaamattomat keskivirheet. Tulosten analysointi perustuu korjattuihin keskivirheisiin.

Tulosten tulkinnassa on huomioitava, että tutkielman empiiriset testit kuvaavat ensisijaisesti tuloerojen ja terveyden välistä korrelaatiota, eikä niiden perusteella voida välttämättä suoraan tehdä päätelmiä tuloerojen ja terveyden välillä vallitsevista kausaalisuussuhteista.

Taulukko 8. Pitkäaikaissairastavuuden todennäköisyyden probit-mallin marginaalivaikutukset vuonna 2003 (tuloerot laskettu maakunnittain)

2003 Muuttuja	Malli 2			Malli 3		
	Marginaali- vaikutus	Korjattu keskivirhe	Korjaamaton keskivirhe	Marginaali- vaikutus	Korjattu keskivirhe	Korjaamaton keskivirhe
Gini-kerroin1	0,285	0,313	0,331	-0,239	0,687	0,687
Keskimääräinen tulo1	-0,772	0,485	0,459	-0,884	0,503*	0,477
R				-0,028	0,021	0,022
QR				0,089	0,073	0,078
Tulot	-0,065	0,034*	0,051	-0,015	0,040	0,060
Tulot <sup>2</sup>	0,004	0,005	0,007	-0,001	0,006	0,008
Ikä	0,019	0,002***	0,002	0,020	0,002***	0,002
Ikä <sup>2</sup>	0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli	-0,043	0,011***	0,010	-0,043	0,011***	0,010
Siviilisääty	-0,029	0,008***	0,012	-0,027	0,007***	0,012
Keskiaste	-0,014	0,013	0,013	-0,014	0,013	0,013
Alin korkea-aste	-0,037	0,016**	0,017	-0,035	0,016**	0,017
Korkea-aste	-0,094	0,013***	0,018	-0,092	0,013***	0,018
Ylempi toimihenkilö	0,022	0,025	0,023	0,022	0,026	0,023
Alempi toimihenkilö	-0,017	0,022	0,020	-0,016	0,022	0,020
Yrittäjä	-0,028	0,016*	0,020	-0,030	0,017*	0,020
Maatalousyrittäjä	0,049	0,021**	0,024	0,047	0,021**	0,024
Opiskelija	0,071	0,026***	0,028	0,068	0,026***	0,028
Eläkeläinen	0,273	0,020***	0,022	0,271	0,020***	0,022
Pitkäaikaistyötön	0,085	0,029***	0,025	0,078	0,031**	0,025
Muu	0,054	0,026**	0,032	0,049	0,027*	0,032
Kotitalouden koko	-0,025	0,005***	0,005	-0,027	0,005***	0,005
Havaintojen lukumäärä	10829			10829		

Koulutus-muuttujassa vertailuryhmänä ovat perusasteen koulutuksen suorittaneet henkilöt ja sosioekonominen asema –muuttujassa työntekijät. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.

Taulukko 9. Pitkäaikaissairastavuuden todennäköisyyden probit-mallin marginaalivaikutukset vuonna 2004 (tuloerot laskettu maakunnittain)

2004 Muuttuja	Malli 2			Malli 3		
	Marginaali- vaikutus	Korjattu keskivirhe	Korjaamaton keskivirhe	Marginaali- vaikutus	Korjattu keskivirhe	Korjaamaton keskivirhe
Gini-kerroin1	0,010	0,305	0,258	0,270	0,474	0,499
Keskimääräinen tulo1	-0,023	0,476	0,398	-0,081	0,513	0,412
R				0,009	0,016	0,018
QR				-0,035	0,056	0,062
Tulot	-0,109	0,093	0,061	-0,094	0,071	0,071
Tulot <sup>2</sup>	0,028	0,018	0,018	0,025	0,013**	0,018
Ikä	0,018	0,002***	0,002	0,018	0,002***	0,002
Ikä <sup>2</sup>	0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli	-0,018	0,012	0,011	-0,018	0,012	0,011
Siviilisääty	-0,009	0,011	0,013	-0,009	0,011	0,013
Keskiaste	-0,021	0,011*	0,013	-0,021	0,011*	0,013
Alin korkea-aste	-0,048	0,017***	0,018	-0,048	0,017***	0,018
Korkea-aste	-0,082	0,014***	0,019	-0,082	0,014***	0,019
Ylempi toimihenkilö	0,014	0,020	0,023	0,015	0,021	0,023
Alempi toimihenkilö	-0,004	0,019	0,020	-0,004	0,019	0,020
Yrittäjä	-0,030	0,016*	0,021	-0,030	0,016*	0,021
Maatalousyrittäjä	0,058	0,026**	0,026	0,058	0,026**	0,026
Opiskelija	0,070	0,032**	0,026	0,069	0,031**	0,026
Eläkeläinen	0,313	0,014***	0,022	0,313	0,014***	0,022
Pitkäaikaistyötön	0,107	0,018***	0,024	0,106	0,018***	0,024
Muu	0,054	0,026**	0,030	0,053	0,026**	0,030
Kotitalouden koko	-0,018	0,005***	0,004	-0,018	0,006***	0,005
Havaintojen lukumäärä	10911			10911		

Koulutus-muuttujassa vertailuryhmänä ovat perusasteen koulutuksen suorittaneet henkilöt ja sosioekonominen asema –muuttujassa työntekijät. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.



Taulukoissa 6 ja 7 esitettävien probit-mallin parametriestimaateista voidaan havaita, että mallissa 1 Gini-kertoimen kerroinestimaatti on negatiivinen ja tilastollisesti merkitsevä. Kun malliin lisätään kontrollimuuttujia, Gini-kertoimen kerroinestimaatti muuttuu positiiviseksi (malli 2). Kerroinestimaatti jää kuitenkin vaille tilastollista merkitsevyyttä, joten maakunnan tuloeroilla ja pitkäaikaissairastavuudella ei tuloksen mukaan ole tilastollisesti merkitsevää yhteyttä, kun tuloeroja mitataan Gini-kertoimella. Tulokset eivät näin ollen tue tulojen eriarvoisuuden hypoteesia vahvaa versiota (ks. luku 3.2, yhtälö 17). Myös taulukoista 8 ja 9 Gini-kertoimen marginaalivaikutus jää vaille tilastollista merkitsevyyttä. Kerroin on hypoteesin mukaisesti positiivinen, ja Gini-kertoimen kasvu 0,10 yksiköllä lisää pitkäaikaissairastavuuden todennäköisyyttä 2,85 ja 0,10 prosenttiyksikköä vuosina 2003 ja 2004. Koska kertoimet eivät ole tilastollisesti merkitseviä, eivät tulokset tue tulojen eriarvoisuuden hypoteesin vahvaa versiota.<sup>24</sup>

Kun Gini-kerroin korvattiin mallissa 2 Atkinsonin ja Theilin indeksillä, tulokset säilyivät taulukoihin 6 ja 7 nähden muuttumattomina lukuun ottamatta Theilin indeksin kerrointa vuonna 2003 (ks. liite 3, taulukko 1). Eräs mahdollinen selitys poikkeavalle tulokselle voi löytyä Theilin indeksin suuremmasta vaihtelusta Gini-kertoimeen verrattuna (ks. liite 2, taulukko 2). Theilin indeksi voi tätä kautta saada suuremman selitysvoiman, ja kertoimesta tulee mallissa tilastollisesti merkitsevä. Tutkimuksissa tulokset tuloerojen ja terveyden välisestä yhteydestä eivät yleensä ole muuttuneet, vaikka tuloerojen mittaamisessa olisikin käytetty useita eri eriarvoisuusmittoja (ks. esim. Gerdtham & Johannesson 2004, 238).

Tulokset Gini-kertoimen ja pitkäaikaissairastavuuden välisestä yhteydestä muuttuvat jonkin verran, kun tuloerot lasketaan seutukunnittain (ks. liite 4, taulukot 1 ja 2). Gini-kertoimen kerroinestimaatit ovat tällöinkin positiivisia, mutta taulukosta 6 poiketen kerroin on tilastollisesti merkitsevä vuonna 2003. Poikkeava tulos voi selittyä seutukuntien Gini-kertoimen suuremmalla vaihtelulla maakuntiin verrattuna (ks. liite 2, taulukko 2), mitä kautta Gini-kerroin saa mallissa suuremman selitysvoiman. Tulos ei näin ollen vastaisi esitettyä väitettä siitä, että maantieteellisesti pienille alueille lasketut tuloerot eivät heijastaisi riittävästi yhteiskunnallista eriarvoisuutta eikä tuloerojen yhteyttä terveyteen tästä syystä esiintyisi (ks. Subramanian & Kawachi 2004, 82). Tulokinnassa on kuitenkin otettava huomioon, että

---

<sup>24</sup> Tulokset muuttuivat jonkin verran, kun Gini-kertoimessa, keskimääräisessä tulossa sekä tulot-muuttujassa käytettiin ekvivalenssiskaalana OECD-skaalaa ja modifioitua OECD-skaalaa. Tällöin tuloerojen ja pitkäaikaissairastavuuden välille saatiin tilastollisesti merkitsevä positiivinen yhteys vuonna 2003.

seutukuntien tuloerojen ja pitkäaikaissairastavuuden välillä ei esiinny yhteyttä vuoden 2004 aineistossa.

Yhteenvedona edellä esitetyistä tuloksista voidaan todeta, että maakunnan tuloeroilla ja pitkäaikaissairastavuudella ei ollut yhteyttä Suomessa vuosina 2003 ja 2004, kun tuloeroja mitattiin Gini-kertoimella. Vuoden 2003 osalta tulokset muuttuivat, kun tuloerot laskettiin seutukunnittain. Tällöin Gini-kertoimen ja pitkäaikaissairastavuuden välille saatiin tilastollisesti merkitsevä positiivinen yhteys vuonna 2003, ja tulokset tukivat esitettyä hypoteesia tuloerojen ja terveyden välisestä yhteydestä. Kun maakunnan Gini-kerroin korvattiin Theilin indeksillä, olivat tulokset tuloerojen ja pitkäaikaissairastavuuden yhteydestä vuoden 2003 osalta samansuuntaisia kuin seutukuntakohtaisessa tarkastelussa.

Tutkielmassa testataan suhteellisen tulon hypoteesia (ks. luku 3.1) alueen keskimääräinen tulo-muuttujan avulla. Taulukoiden 6 ja 7 mallista 2 havaitaan, että maakunnan keskimääräisen tulon kerroin on negatiivinen. Negatiivinen kerroin tarkoittaa, että keskimääräisen tulon kasvaessa pitkäaikaissairastavuuden todennäköisyys pienenee. Suhteellisen tulon hypoteesin mukaan terveys heikkenee, kun yksilö on taloudellisesti huono-osainen suhteessa viiteryhmään. Alueen keskimääräisen tulon kasvaminen merkitsisi hypoteesin mukaan yksilön pitkäaikaissairastavuuden todennäköisyyden kasvamista. Parametriestimaatti antaa näin ollen hypoteesin vastaisen tuloksen. Estimaatti jää kuitenkin vaille tilastollista merkitsevyyttä, joten maakunnan keskimääräisellä tulolla ja pitkäaikaissairastavuudella ei tuloksen mukaan ole yhteyttä. Myös seutukunnittain lasketun keskimääräisen tulon parametriestimaatti osoittautuu negatiiviseksi, mutta jää vaille tilastollista merkitsevyyttä (ks. liite 4, taulukot 1 ja 2).

Absoluuttisen tulon hypoteesia testataan puolestaan tulot-muuttujan avulla. Vuoden 2003 aineistossa taulukossa 6 tilastollisesti merkitsevä negatiivinen kerroin tukee esitettyä absoluuttisen tulon hypoteesia, jonka mukaan yksilön terveys on sitä parempi, mitä suuremmat tulot hänellä on. Tulojen kasvaessa henkilön pitkäaikaissairastavuuden todennäköisyys pienenee, eli varakkaammat henkilöt ovat pienituloisia terveempiä. Tulos poikkeaa edellä taulukosta 4 tehdystä havainnosta, jonka perusteella tulot olivat keskimäärin suuremmat pitkäaikaissairaiden ryhmässä. Taulukosta 8 nähdään, että tulojen kasvaessa tuhannella eurolla, pitkäaikaissairastavuuden todennäköisyys pienenee 0,065 prosenttiyksikköä. Koska tulojen neliömuuttujan jää vaille tilastollista merkitsevyyttä, tulojen



ja terveyden välinen yhteys ei ole konkaavi (ks. taulukko 6). Tulos on tulojen ja terveyden välisen yhteyden konkaavisuusolettaman vastainen (vrt. luku 3.2, kuvio 5). Vuonna 2004 tulojen ja pitkäaikaissairastavuuden välillä on negatiivinen yhteys, joka jää kuitenkin vaille tilastollista merkitsevyyttä (ks. taulukko 7).

Taulukoiden 6 ja 7 mallissa 2 muiden yksilö- ja kotitalouden piirteitä kuvaavien muuttujien kerroinestimaatit ovat pitkälti ennako-odotusten mukaisia. Ikä-muuttujan positiivinen kerroinestimaatti osoittaa, että iän myötä pitkäaikaissairastavuuden todennäköisyys kasvaa. Tulos oli ennakoitavissa jo taulukon 4 tuloksista. Vuonna 2003 yksi lisävuosi kasvatti pitkäaikaissairastavuuden todennäköisyyttä 1,9 prosenttiyksikköä ja vuonna 2004 1,8 prosenttiyksikköä (ks. taulukot 8 ja 9). Iän neliö -muuttujan parametristimaatti saa tilastollisesti merkitsevän arvon nolla, joten yksilön iän ja pitkäaikaissairastavuuden välinen yhteys on lineaarinen (ks. taulukot 6 ja 7).

Tutkimustulokset osoittavat myös, että naisilla pitkäaikaissairastavuuden todennäköisyys on hieman suurempi kuin miehillä. Tulos on ennakoitavissa aineiston kuvailevasta analyysistä (ks. liite 2, taulukot 2 ja 3) sekä terveiden ja pitkäaikaissairaiden ryhmien vertailusta taulukossa 4. Taulukosta 8 voidaan havaita, että pitkäaikaissairastavuuden todennäköisyys on miehillä 4,3 prosenttiyksikköä pienempi kuin naisilla vuonna 2003. Vuonna 2004 marginaalivaikutus on samansuuntainen, mutta jää vaille tilastollista merkitsevyyttä. Yksi selitys naisten suuremmalle pitkäaikaissairastavuudelle voi löytyä tutkielman haastatteluun perustuvasta aineistosta. On mahdollista, että naiset raportoivat miehiä herkemmin omista sairauksistaan ja naisten miehiä suurempi pitkäaikaissairastavuuden todennäköisyys selittyisi tätä kautta. Toisaalta naisten suurempi pitkäaikaissairastavuus voi olla seurausta naisten miehiä korkeammasta elinajanodotteesta ja sitä kautta lisääntyvistä pitkäaikaissairauksista.

Odotusten mukaisesti tuloksista ilmenee myös, että naimisissa olevilla henkilöillä pitkäaikaissairastavuuden todennäköisyys on muita siviilisäätyryhmiä pienempi. Myös perheen koon kasvaminen pienentää pitkäaikaissairastavuuden todennäköisyyttä. Kun tarkastellaan koulutuksen vaikutusta pitkäaikaissairastavuuteen, voidaan taulukoista 7 ja 8 havaita koulutuksen terveyttä suojaava vaikutus. Taulukoiden 8 ja 9 mallin 2 tuloksista nähdään, että vuonna 2003 pitkäaikaissairastavuuden todennäköisyys oli alimman korkea-asteen koulutuksen suorittaneilla 3,7 prosenttiyksikköä ja korkea-asteen koulutuksen suorittaneilla 9,4 prosenttiyksikköä pienempi kuin perusasteen koulutuksen suorittaneilla.

Keskiasteen koulutuksen suorittamisella ei ollut tällöin tilastollisesti merkitsevää yhteyttä pitkäaikaissairastavuuteen. Koulutuksen pitkäaikaissairauksilta suojaava vaikutus oli ennakoitavissa jo taulukon 4 tuloksista, joiden perusteella pitkäaikaissairaat olivat keskimäärin vähemmän koulutettuja kuin terveet. Sosioekonomisen aseman osalta tuloksista havaitaan puolestaan, että ainoastaan yrittäjillä pitkäaikaissairastavuuden todennäköisyys on tilastollisesti merkitsevästi pienempi kuin työntekijäryhmällä. Työntekijöihin verrattuna suurin pitkäaikaissairauksien todennäköisyys on odotusten mukaisesti eläkeläisillä ja pitkäaikaistyöttömillä.

Tulojen eriarvoisuuden hypoteesin heikkoa versiota testataan mallilla 3, jonka tulokset esitetään taulukoissa 6-9. Taulukoista voidaan havaita, että *QR*-muuttujan parametriestimaatti on tulojen eriarvoisuuden hypoteesin heikon version mukaisesti negatiivinen vuonna 2004. Vuonna 2003 estimaatti saa hypoteesin vastaisen positiivisen kertoimen. *QR*-muuttujan parametriestimaatti jää kuitenkin vaille tilastollista merkitsevyyttä sekä vuonna 2003 että vuonna 2004. Yksilön tulotasosta riippumatta tuloeroilla ei ole näin ollen yhteyttä pitkäaikaissairastavuuteen, kun maakunnan eriarvoisuutta mitataan Gini-kertoimella. Tulokset eivät tue tulojen eriarvoisuuden hypoteesin heikkoa versiota.

Tutkielmassa selvitetään myös, muuttuuko tuloerojen ja pitkäaikaissairastavuuden välinen yhteys, mikäli samanaikaisesti otetaan huomioon alueen köyhyys sekä tarkastellaan, minkälainen yhteys maakunnittain lasketulla köyhyydellä on pitkäaikaissairastavuuteen. Tätä selvitetään laajentamalla mallia 2 lisäämällä siihen muuttujiksi maakunnan köyhyyttä kuvaavat päälukumitta ja köyhyyskuilu. Estimointitulokset pitkäaikaissairastavuuden osalta esitetään liitteessä 5. Tuloksista voidaan nähdä, että päälukumitta ja köyhyyskuilu saavat negatiivisen kertoimen. Kertoimet jäävät kuitenkin vaille tilastollista merkitsevyyttä, joten pitkäaikaissairastavuudella ja alueellisella köyhyydelle ei siten näyttäisi olevan yhteyttä. Hypoteesia köyhyyden terveyttä huonontavasta vaikutuksesta ei voida kuitenkaan täysin luotettavasti kumota ja osoittaa vaikutussuunnaltaan päinvastaiseksi, sillä päälukumitan ja köyhyyskuilun negatiivinen kerroin jää tuloksissa vaille tilastollista merkitsevyyttä. Esimerkiksi taulukosta 4 havaittiin, että vuonna 2004 maakunnan köyhyysaste oli tilastollisesti merkitsevästi keskimäärin pienempi terveiden ryhmässä pitkäaikaissairaisiin verrattuna.



Köyhyysmuuttujien lisääminen malliin 2 ei muuta tuloksia tuloerojen ja pitkäaikaissairastavuuden välisestä yhteydestä vuonna 2004. Vuoden 2003 osalta Gini-kertoimen ja pitkäaikaissairastavuuden välille saadaan tilastollisesti merkitsevä positiivinen kerroin, kun malliin 2 lisätään maakunnan köyhyyttä kuvaava päälukumitta. Edellä esitettyyn nähden poikkeava tulos voi olla seurausta multikollineaarisuudesta selittävien muuttujien kesken. Multikollineaarisuusongelmassa selittäjät korreloivat voimakkaasti keskenään, jolloin regressiokertoimien estimaateista tulee epävakaita ja keskivirheet kasvavat. Vuoden 2003 aineistossa esimerkiksi päälukumitan ja maakunnan keskimääräisen tulon välinen korrelaatio oli -0,760. Kun malliin 2 lisättiin maakunnan päälukumitta, keskimääräinen tulo-muuttujan keskivirhe kasvoi arvosta 1,266 arvoon 2,832. Gini-kertoimen keskivirhe nousi vastaavasti arvosta 0,821 arvoon 1,380, ja kerroin muuttui tilastollisesti merkitseväksi.

## 6.2 Tuloerojen yhteys itse arvioituun yleiseen terveydentilaan

Taulukoissa 10 ja 11 esitetään järjestetyn probit -mallin estimointitulokset tuloerojen ja yleisen terveydentilan välistä yhteyttä tarkasteleville yhtälöille sekä marginaalivaikutukset selitettävän muuttujan alimman (yleinen terveydentila=0) ja ylimmän (yleinen terveydentila=3) vastauskategorian osalta. Kuten edellä pitkäaikaissairastavuuden kohdalla testataan mallin 2 avulla tulojen eriarvoisuuden hypoteesin vahvaa versiota ja mallilla 3 hypoteesin heikkoa versiota.

Järjestetty probit-mallin peruskertoimien etumerkeistä voidaan päätellä, mihin suuntaan yleinen terveydentila muuttuu selittäjässä tapahtuvan muutoksen seurauksena. Kertoimen ollessa positiivinen yleinen terveydentila paranee ja päinvastoin. Mikäli tulokset tukevat tulojen eriarvoisuuden hypoteesin vahvaa versiota, tuloeroja kuvaavan muuttujan kerroinestimaatti on negatiivinen ja tilastollisesti merkitsevä. Vastaavasti *QR*-muuttujan tilastollisesti merkitsevä positiivinen kerroin tukee hypoteesin heikkoa versiota. Kertoimen avulla ei voida kuitenkaan tehdä johtopäätöksiä muutoksen suuruudesta. Taulukoissa 10 ja 11 esitetyt marginaalivaikutukset kuvaavat, kuinka paljon yleisen terveydentilan alimman ja ylimmän vastauskategorian todennäköisyys muuttuu selittäjässä tapahtuvan muutoksen seurauksena, kun muut selittävät muuttujat ovat keskiarvossaan. Luvussa 5.4.2 havaittiin, että

Taulukko 10. Yleisen terveydentilan todennäköisyyden järjestetty probit -mallin kerroinestimaatit ja marginaalivaikutukset vuonna 2003 (tuloerot laskettu maakunnittain)

2003 Muuttuja	Malli 2			Yleinen terveydentila=0			Yleinen terveydentila=3		
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe
Gini-kerroin1	-0,958	0,492*	0,729	0,108	0,056*	0,082	-0,377	0,193*	0,287
Keskimäär. tulo1	3,747	0,794***	1,017	-0,421	0,096***	0,115	1,475	0,310***	0,400
Tulot	0,433	0,109***	0,117	-0,049	0,011***	0,013	0,171	0,043***	0,046
Tulot <sup>2</sup>	-0,037	0,017**	0,015	0,004	0,002**	0,002	-0,015	0,007**	0,006
Ikä	-0,059	0,003***	0,005	0,007	0,000***	0,001	-0,023	0,001***	0,002
Ikä <sup>2</sup>	0,000	0,000***	0,000	0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli	-0,043	0,026	0,023	0,005	0,003*	0,003	-0,017	0,010	0,009
Siviilisaäty	0,083	0,025***	0,027	-0,009	0,003***	0,003	0,032	0,010***	0,011
Keskiaste	0,112	0,034***	0,028	-0,012	0,004***	0,003	0,044	0,014***	0,011
Alin korkea-aste	0,232	0,034***	0,039	-0,023	0,003***	0,003	0,092	0,013***	0,015
Korkea-aste	0,416	0,034***	0,043	-0,037	0,003***	0,003	0,165	0,013***	0,017
Ylempi toimihlö	0,092	0,042**	0,051	-0,010	0,004**	0,005	0,036	0,017**	0,020
Alempi toimihlö	0,113	0,036***	0,045	-0,012	0,004***	0,004	0,045	0,014***	0,018
Yrittäjä	-0,011	0,049	0,044	0,001	0,006	0,005	-0,004	0,019	0,017
Maatalousyrittäjä	-0,213	0,043***	0,051	0,028	0,006***	0,008	-0,082	0,016***	0,019
Opiskelija	-0,012	0,044	0,062	0,001	0,005	0,007	-0,005	0,017	0,025
Eläkeläinen	-0,454	0,052***	0,048	0,062	0,008***	0,008	-0,173	0,019***	0,017
Pitkäaikaistyötn	-0,160	0,059***	0,054	0,020	0,009**	0,008	-0,062	0,023***	0,021
Muu	-0,152	0,062**	0,069	0,019	0,009**	0,010	-0,059	0,024**	0,026
Kotitalouden koko	0,041	0,008***	0,010	-0,005	0,001***	0,001	0,016	0,003***	0,004
Havaintojen lkm	10826			10826			10826		
Pseudo R <sup>2</sup>	0,106								

Koulutus-muuttujassa vertailuryhmänä ovat perusasteen koulutuksen suorittaneet henkilöt ja sosioekonominen asema –muuttujassa työntekijät. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevää 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.

Taulukko 11. Yleisen terveydentilan todennäköisyyden järjestetty probit -mallin kerroinestimaatit ja marginaalivaikutukset vuonna 2004 (tuloerot laskettu maakunnittain)

2004 Muuttuja	Malli 2			Yleinen terveydentila=0			Yleinen terveydentila=3		
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe
Gini-kerroin1	-0,031	0,706	0,555	0,004	0,080	0,063	-0,012	0,277	0,218
Keskimäär. tulo1	1,738	0,809**	0,865	-0,196	0,088**	0,098	0,683	0,317**	0,340
Tulot	0,499	0,167***	0,112	-0,056	0,020***	0,013	0,196	0,066***	0,044
Tulot <sup>2</sup>	-0,025	0,009***	0,014	0,003	0,001**	0,002	-0,010	0,004***	0,005
Ikä	-0,061	0,003***	0,004	0,007	0,000***	0,001	-0,024	0,001***	0,002
Ikä <sup>2</sup>	0,000	0,000***	0,000	0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli	-0,049	0,029*	0,023	0,006	0,003*	0,003	-0,019	0,011*	0,009
Siviilisaäty	0,097	0,021***	0,027	-0,011	0,002***	0,003	0,038	0,008***	0,011
Keskiaste	0,106	0,023***	0,028	-0,012	0,003***	0,003	0,042	0,009***	0,011
Alin korkea-aste	0,242	0,035***	0,040	-0,024	0,003***	0,003	0,096	0,034***	0,016
Korkea-aste	0,358	0,038***	0,043	-0,033	0,002***	0,003	0,142	0,015***	0,017
Ylempi toimihlö	0,097	0,032***	0,052	-0,010	0,003***	0,005	0,038	0,013***	0,021
Alempi toimihlö	0,139	0,031***	0,044	-0,014	0,003***	0,004	0,055	0,012***	0,018
Yrittäjä	0,022	0,036	0,046	-0,002	0,004	0,005	0,009	0,014	0,018
Maatalousyrittäjä	-0,172	0,049***	0,055	0,022	0,007***	0,008	-0,066	0,019***	0,021
Opiskelija	0,028	0,064	0,058	-0,003	0,007	0,006	0,011	0,025	0,023
Eläkeläinen	-0,481	0,045***	0,048	0,067	0,006***	0,008	-0,182	0,017***	0,017
Pitkäaikaistyötn	-0,248	0,027***	0,052	0,033	0,004***	0,008	-0,095	0,010***	0,019
Muu	-0,179	0,082**	0,065	0,023	0,011**	0,009	-0,069	0,031**	0,024
Kotitalouden koko	0,038	0,007***	0,010	-0,004	0,001***	0,001	0,015	0,003***	0,004
Havaintojen lkm	10904			10904			10904		
Pseudo R <sup>2</sup>	0,113								

Koulutus-muuttujassa vertailuryhmänä ovat perusasteen koulutuksen suorittaneet henkilöt ja sosioekonominen asema –muuttujassa työntekijät. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevää 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.



järjestetyn probit-mallin tapauksessa peruserroin määrittää marginaalivaikutuksen suunnan yksiselitteisesti ainoastaan alimman ja ylimmän vastauskategorian kohdalla. Tästä johtuen marginaalivaikutukset raportoidaan taulukoissa 10 ja 11 ainoastaan edellä mainittujen vastauskategorioiden osalta. Marginaalivaikutukset voidaan tulkita prosenttiyksiköinä, kun parametriestimaatit kerrotaan luvulla sata. Gini-kertoimessa, alueen keskimääräisessä tulossa ja yksilöiden käytettävissä olevissa tuloissa käytetyn skaalan vuoksi marginaalivaikutusten tulkinnassa sovelletaan samaa muunnostapaa kuin edellä probit-mallin kohdalla (vrt. luku 6.1). Lisäksi taulukoissa ilmoitetaan jokaiselle selittäjälle sekä klusteroinnin huomioivat että korjaamattomat keskivirheet. Tulosten analysointi perustuu korjattuihin keskivirheisiin.

Maakunnan Gini-kertoimelle ja yleiselle terveydentilalle saadaan taulukossa 10 tilastollisesti merkitsevä negatiivinen yhteys. Tuloerojen kasvaessa yleinen terveydentila heikkenee. Tulos tukee näin ollen tulojen eriarvoisuuden hypoteesin vahvaa versiota vuonna 2003. Taulukon 10 viidennessä ja kahdeksannesta sarakkeesta nähdään, että alimman ja ylimmän vastauskategorian marginaalivaikutukset ovat esitetyn teorian mukaisesti erimerkkiset: Gini-kertoimen kasvaessa 0,10 yksikköä huonon tai melko huonon yleisen terveydentilan todennäköisyys nousee 1,08 prosenttiyksikköä ja hyvän terveydentilan todennäköisyys laskee 3,77 prosenttiyksikköä. Vuoden 2004 osalta maakunnan tuloerojen ja yleisen terveydentilan välille ei saada tilastollisesti merkitsevää yhteyttä, kun eriarvoisuutta mitataan Gini-kertoimella (ks. taulukko 11).<sup>25</sup>

Tulokset tuloerojen ja yleisen terveydentilan välisestä yhteydestä säilyivät pääpiirteissään muuttumattomina, vaikka Gini-kerroin korvattiin maakuntakohtaisella Atkinsonin ja Theilin indeksillä (ks. liite 6). Vuoden 2003 osalta eriarvoisuusmitat saivat hypoteesin mukaisen negatiivisen ja tilastollisesti merkitsevän kertoimen. Atkinsonin ja Theilin indeksin tilastollinen merkitsevyys oli tällöin jonkin verran suurempi kuin Gini-kertoimella taulukossa 10. Vuonna 2004 tuloeroilla ja yleisellä terveydentilalla ei ollut tilastollisesti merkitsevää yhteyttä. Theilin indeksi sai tällöin hypoteesin vastaisen positiivisen kertoimen, joka ei kuitenkaan eronnut tilastollisesti merkitsevästi nolasta.

---

<sup>25</sup> Estimointitulokset tuloerojen ja yleisen terveydentilan välisestä yhteydestä eivät muuttuneet, vaikka Gini-kertoimessa, keskimääräisessä tulossa sekä tulot-muuttujassa käytettiin ekvivalenssiskaalana per capitaan sijaan sekä OECD-skaalaa että modifioitua OECD-skaalaa.

Tulokset muuttuivat hieman, kun tuloerot laskettiin seutukunnittain (ks. liite 7). Gini-kerroin sai negatiivisen kertoimen, mutta jäi maakuntakohtaisesta tarkastelusta poiketen vaille tilastollista merkitsevyyttä vuonna 2003. Vuonna 2004 Gini-kertoimen parametriestimaatti sai yllättäen hypoteesin vastaisen positiivisen kertoimen, mutta ei myöskään tällöin eronnut nolasta tilastollisesti merkitsevästi. Itse arvioidulla yleisellä terveydentilalla ja tuloeroilla ei siten ollut yhteyttä vuosina 2003 ja 2004, kun eriarvoisuutta mitattiin seutukunnittain.

Tulosten muutosherkkyyttä tarkasteltiin myös korvaamalla selitettävänä muuttujana oleva yleinen terveydentila kutakin terveydentilan luokkaa vastaavalla 15D-terveysmittarin keskimääräisellä arvolla (ks. s. 49 ja liitteen 9 taulukko 1). Liitteen 9 taulukossa 2 esitetyt estimointitulokset vastaavat maakunnan Gini-kertoimen ja yleisen terveydentilan välillä edellä tehtyjä havaintoja. Terveysteen liittyvää elämänlaatua mittaavan 15D-muuttujan ja Gini-kertoimen välille saadaan tilastollisesti merkitsevä negatiivinen yhteys vuonna 2003, kun taas vuotta myöhemmin yhteyttä ei esiinny. Tulokset muuttuvat yleiseen terveydentilaan verrattuna, kun tuloerot lasketaan seutukunnittain. Seutukunnan Gini-kerroin saa tilastollisesti merkitsevän negatiivisen kertoimen vuonna 2003. Vuotta myöhemmin yhteyttä ei myöskään tässä tapauksessa esiinny. Yhteenvetona voidaan todeta, että sekä maakunnan että seutukunnan sisäisten tuloerojen kasvu pienensi terveyteen liittyvää elämänlaatua vuonna 2003, kun sitä mitattiin 15D-terveysmittarilla. Tulokset tukevat tulojen eriarvoisuuden hypoteesia tällöin.

Maakunnan keskimääräiselle tulolle ja yleiselle terveydentilalle saadaan taulukoissa 10 ja 11 tilastollisesti merkitsevä positiivinen yhteys. Myös seutukunnan keskimääräisen tulon ja yleisen terveydentilan välillä on tilastollisesti merkitsevä positiivinen yhteys (ks. liite 7). Alueen keskimääräisen tulon kasvaessa yleinen terveydentila paranee. Saatu tulos on suhteellisen tulon hypoteesin vastainen (vrt. luku 3.1). Toisaalta on esitetty myös, että yhteisön korkea keskimääräinen tulotaso voi olla terveydentilaa suojaava. Millerin ja Paxsonin (2000, 5) mukaan varakkaiden henkilöiden keskittyminen tietylle alueelle voi johtaa esimerkiksi julkishyödykkeisiin käytettävien varojen kasvuun ja ympäristöä suojelevien säännösten suurempaan kysyntään, jotka parantavat siten kaikkien alueella asuvien terveyttä.

Myös yksilön tuloilla on tulosten mukaan yleistä terveydentilaa parantava vaikutus. Tulos oli ennakoitavissa jo edellä hyvän sekä huonon tai melko huonon terveydentilan omaavien ryhmien vertailusta taulukossa 5. Taulukoiden 10 ja 11 tulot-muuttujan tilastollisesti



merkitsevä positiivinen kerroin vahvistaa tuloksen. Taulukoista nähdään, että huonon tai melko huonon yleisen terveydentilan todennäköisyys pieneni 0,049 ja 0,056 prosenttiyksikköä vuosina 2003 ja 2004, kun tulot kasvoivat tuhat euroa. Hyvän yleisen terveydentilan todennäköisyys kasvoi vastaavasti 0,177 ja 0,196 prosenttiyksikköä kyseessä olevina vuosina. Tulos on yhtenevä absoluuttisen tulon hypoteesin kanssa. Tulojen ja yleisen terveydentilan välinen yhteys ei kuitenkaan ole lineaarinen, vaan pienituloisilla tulojen kasvu vaikuttaa terveydentilaan enemmän kuin suurituloisilla. Tämä voidaan havaita tulojen neliö-muuttujan negatiivisesta parametriestimaatista taulukoissa 10 ja 11. Tulos on yhtenevä tulojen ja terveyden välisen yhteyden konkaavisuusolettaman (ks. luku 3.2, kuvio 5) kanssa.

Yleisen terveydentilan ja henkilön iän välille saadaan taulukoissa 10 ja 11 tilastollisesti merkitsevä negatiivinen yhteys. Odotusten mukaisesti iän kasvaessa yleinen terveydentila heikkenee. Vaikutus huonoon tai melko huonoon terveydentilaan ei tosin ole merkittävän suuri. Vuosien 2003 ja 2004 aineistoissa yksi lisävuosi kasvatti huonon tai melko huonon yleisen terveydentilan todennäköisyyttä 0,7 prosenttiyksikköä. Vastaavasti hyvän yleisen terveydentilan todennäköisyys pieneni 2,3 ja 2,4 prosenttiyksikköä tarkasteluajankohtina. Kuten edellä pitkäaikaissairauksien kohdalla myös yleisen terveydentilan ja yksilön iän välinen yhteys on lineaarinen. Tämä voidaan päätellä iän neliö-muuttujan tilastollisesti merkitsevästä, arvon nolla saavasta parametriestimaatista taulukoissa 10 ja 11.

Tulokset vahvistavat aineiston kuvailevasta analyysistä tehtävän havainnon, jonka mukaan naiset kokevat yleisen terveydentilansa miehiä paremmaksi (ks. liite 2, taulukot 2 ja 3). Vastaava tulos oli nähtävissä myös hyvän sekä huonon tai melko huonon terveydentilan omaavien henkilöiden vertailusta (ks. taulukko 5). Taulukoissa 10 ja 11 sukupuoli-muuttuja saa negatiivisen kertoimen, joka merkitsee, että naisilla yleinen terveydentila on miehiä parempi. Kerroin jää vaille tilastollista merkitsevyyttä vuoden 2003 aineistossa, mutta huonon tai melko huonon terveydentilan todennäköisyyden muutosta kuvaava marginaalivaikutus eroaa tällöin nolasta tilastollisesti merkitsevästi. Sen mukaan huonon tai melko huonon terveydentilan todennäköisyys on miehillä 0,5 prosenttiyksikköä korkeampi kuin naisilla. Taulukosta 11 nähdään puolestaan, että vuonna 2004 miehillä huonon tai melko huonon yleisen terveydentilan todennäköisyys oli 0,6 prosenttiyksikköä suurempi ja hyvän yleisen terveydentilan todennäköisyys 1,9 prosenttiyksikköä pienempi kuin naisilla. Vaikka tutkielman tulosten perusteella naisilla on miehiä enemmän pitkäaikaissairauksia, ei se näytä

vaikuttavan heidän itse arvioituun yleiseen terveydentilaan, sillä naiset raportoivat yleisen terveydentilansa miehiä paremmaksi.

Siviilisäädylle ja kotitalouden koolle saadaan odotusten mukaisesti tilastollisesti merkitsevä positiivinen yhteys (ks. taulukot 10 ja 11). Tulosten mukaan naimisissa olevilla ja suurissa perheissä asuvilla henkilöillä yleinen terveydentila on muita siviilisäätyryhmiä ja pienissä kotitalouksissa asuvia henkilöitä parempi. Tulos on yhdenmukainen edellä pitkäaikaissairastavuuden osalta esitettyjen tulosten kanssa. Ennakko-odotusten mukaisesti ihmissuhteilla ja sosiaalisella verkostolla näyttäisi siten olevan terveyttä suojaava vaikutus. Siviilisäädyn ja kotitalouden koon yhteys terveyteen voi selittyä osaltaan myös henkilöiden iän kautta. Iäkkäät ihmiset ovat usein yksin asuvia ja leskiä sekä sairastavat keskimäärin nuoria enemmän.

Tarkasteltaessa koulutuksen vaikutusta terveydentilaan havaitaan taulukoista 10 ja 11, että kaikki kolme eri koulutusryhmä-muuttujaa saavat tilastollisesti merkitsevän positiivisen kertoimen, kun niitä verrataan perusasteen koulutuksen suorittaneiden ryhmään. Tuloksesta voidaan päätellä, että koulutuksella on yleistä terveydentilaa suojaava vaikutus. Kun tarkastellaan taulukoiden kolmatta ja neljättä saraketta, voidaan niistä havaita koulutuksen nostavan hyvän terveydentilan todennäköisyyttä ja pienentävän huonon tai melko huonon terveydentilan todennäköisyyttä johdonmukaisesti. Esimerkiksi vuonna 2003 keskiasteen koulutuksen suorittaneilla hyvän terveydentilan todennäköisyys oli 4,4 prosenttiyksikköä, alimman korkea-asteen koulutuksen suorittaneilla 9,2 prosenttiyksikköä ja korkea-asteen koulutuksen suorittaneilla jopa 16,5 prosenttiyksikköä korkeampi kuin perusasteen koulutuksen suorittaneilla henkilöillä.

Sosioekonomisen aseman osalta tuloksista nähdään, että työntekijöihin verrattuna ainoastaan ylemmillä ja alemmilla toimihenkilöillä on tilastollisesti merkitsevästi parempi yleinen terveydentila. Muista sosioekonomista ryhmistä ennakko-odotusten mukaisesti eläkeläiset, maatalousyrittäjät ja pitkäaikaistyöttömät kokevat yleisen terveydentilansa muita huonommaksi. Taulukosta 11 nähdään, että eläkeläisillä hyvän terveydentilan todennäköisyys oli jopa 18,2 prosenttiyksikköä pienempi kuin työntekijöillä vuonna 2004. Pitkäaikaistyöttömillä vastaava luku on 9,5 prosenttiyksikköä.



Taulukko 12. Yleisen terveydentilan todennäköisyyden järjestetty probit -mallin kerroinestimaatit ja marginaalivaikutukset vuonna 2003 (tuloerot laskettu maakunnittain)

2003 Muuttuja	Malli 3			Yleinen terveydentila=0			Yleinen terveydentila=3		
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe
Gini-kerroin1	-0,003	1,223	1,513	0,000	0,137	0,170	-0,001	0,482	0,596
Keskimäär. tulo1	4,324	0,785***	1,054	-0,485	0,097***	0,119	1,703	0,307***	0,415
R	0,071	0,044	0,048	-0,008	0,005	0,005	0,028	0,017	0,019
QR	-0,208	0,157	0,174	0,023	0,018	0,019	-0,082	0,062	0,069
Tulot	0,227	0,069***	0,135	-0,025	0,008***	0,015	0,089	0,027***	0,053
Tulot <sup>2</sup>	-0,016	0,016	0,017	0,002	0,002	0,002	-0,006	0,006	0,007
Ikä	-0,060	0,003***	0,005	0,007	0,000***	0,001	-0,024	0,001***	0,002
Ikä <sup>2</sup>	0,000	0,000***	0,000	0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli	-0,045	0,026*	0,023	0,005	0,003*	0,003	-0,018	0,010*	0,009
Siviilisääty	0,077	0,025***	0,027	-0,009	0,003***	0,003	0,030	0,010**	0,011
Keskiaste	0,109	0,035***	0,028	-0,012	0,004***	0,003	0,043	0,014***	0,011
Alin korkea-aste	0,226	0,034***	0,039	-0,022	0,003***	0,003	0,090	0,014***	0,016
Korkea-aste	0,408	0,034***	0,043	-0,037	0,003***	0,003	0,162	0,013***	0,017
Ylempi toimihlö	0,090	0,042**	0,051	-0,010	0,004**	0,005	0,035	0,017**	0,020
Alempi toimihlö	0,109	0,036***	0,045	-0,012	0,004***	0,004	0,043	0,014***	0,018
Yrittäjä	-0,005	0,049	0,045	0,001	0,006	0,005	-0,002	0,019	0,018
Maatalousyrittäjä	-0,208	0,043***	0,051	0,027	0,006***	0,008	-0,080	0,016***	0,019
Opiskelija	-0,004	0,043	0,062	0,000	0,005	0,007	-0,001	0,017	0,025
Eläkeläinen	-0,449	0,052***	0,048	0,0612	0,008***	0,008	-0,171	0,019***	0,017
Pitkäaikaistyötön	-0,136	0,062**	0,054	0,017	0,009*	0,007	-0,053	0,024**	0,021
Muu	-0,135	0,064**	0,070	0,017	0,009*	0,009	-0,052	0,025**	0,027
Kotitalouden koko	0,049	0,009***	0,010	-0,006	0,001***	0,001	0,019	0,003***	0,004
Havaintojen lkm	10826			10826			10826		
Pseudo R <sup>2</sup>	0,107								

Koulutus-muuttujassa vertailuryhmänä ovat perusasteen koulutuksen suorittaneet henkilöt ja sosioekonominen asema -muuttujassa työntekijät. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.

Taulukko 13. Yleisen terveydentilan todennäköisyyden järjestetty probit -mallin kerroinestimaatit ja marginaalivaikutukset vuonna 2004 (tuloerot laskettu maakunnittain)

2004 Muuttuja	Malli 3			Yleinen terveydentila=0			Yleinen terveydentila=3		
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe
Gini-kerroin1	-0,755	0,984	1,085	0,085	0,111	0,122	-0,296	0,387	0,426
Keskimäär. tulo1	2,234	0,781***	0,891	-0,252	0,085***	0,101	0,877	0,306***	0,350
R	-0,006	0,023	0,038	0,001	0,003	0,004	-0,002	0,009	0,015
QR	0,063	0,075	0,136	-0,007	0,009	0,015	0,025	0,029	0,053
Tulot	0,342	0,159**	0,133	-0,039	0,019**	0,015	0,134	0,062**	0,052
Tulot <sup>2</sup>	-0,011	0,010	0,020	0,001	0,001	0,002	-0,004	0,004	0,008
Ikä	-0,062	0,003***	0,004	0,007	0,000***	0,001	-0,024	0,001***	0,002
Ikä <sup>2</sup>	0,000	0,000***	0,000	0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli	-0,051	0,028*	0,023	0,006	0,003*	0,003	-0,020	0,011*	0,009
Siviilisääty	0,095	0,021***	0,027	-0,011	0,003***	0,003	0,037	0,008***	0,011
Keskiaste	0,105	0,023***	0,028	-0,012	0,003***	0,003	0,041	0,009***	0,011
Alin korkea-aste	0,239	0,034***	0,040	-0,023	0,003***	0,003	0,095	0,014***	0,016
Korkea-aste	0,352	0,037***	0,043	-0,033	0,002***	0,003	0,139	0,015***	0,017
Ylempi toimihlö	0,094	0,033***	0,052	-0,010	0,003***	0,005	0,037	0,013***	0,021
Alempi toimihlö	0,137	0,031***	0,044	-0,014	0,003***	0,004	0,054	0,012***	0,018
Yrittäjä	0,028	0,036	0,046	-0,003	0,004	0,005	0,011	0,014	0,018
Maatalousyrittäjä	-0,169	0,049***	0,055	0,021	0,007***	0,008	-0,065	0,019***	0,021
Opiskelija	0,037	0,065	0,058	-0,004	0,007	0,006	0,015	0,026	0,023
Eläkeläinen	-0,474	0,045***	0,048	0,065	0,006***	0,008	-0,179	0,017***	0,017
Pitkäaikaistyötön	-0,231	0,029***	0,053	0,030	0,005***	0,008	-0,088	0,011***	0,019
Muu	-0,163	0,083**	0,065	0,021	0,011*	0,009	-0,063	0,031**	0,025
Kotitalouden koko	0,043	0,008***	0,010	-0,005	0,001***	0,001	0,017	0,003***	0,004
Havaintojen lkm	10904			10904			10904		
Pseudo R <sup>2</sup>	0,113								

Koulutus-muuttujassa vertailuryhmänä ovat perusasteen koulutuksen suorittaneet henkilöt ja sosioekonominen asema -muuttujassa työntekijät. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.

Taulukoissa 12 ja 13 esitetään mallin 3 estimointitulokset, joiden avulla pyritään tekemään päätelmiä tulojen eriarvoisuuden hypoteesin heikosta versiosta. Taulukosta 13 nähdään, että *QR*-muuttujan kerroin on hypoteesin mukaisesti positiivinen vuonna 2004. Estimaatti saa hypoteesin vastaisen negatiivisen kertoimen vuonna 2003. *QR*-muuttujan parametriestimaatti jää kuitenkin vaille tilastollista merkitsevyyttä sekä vuonna 2003 että vuonna 2004. Tulokset osoittavat, että tuloerojen ja yleisen terveydentilan välinen yhteys ei vaihtelee yksilöiden suhteellisen tulotason mukaan. Tulokset eivät tue tulojen eriarvoisuuden hypoteesin heikkoa versiota. Tuloterot eivät siten näyttäisi olevan haitallisimpia tulojen suhteen huono-osaisten terveydelle.

Myös yleisen terveydentilan osalta selvitetään, muuttuuko tuloerojen ja yleisen terveydentilan välinen yhteys, mikäli samanaikaisesti otetaan huomioon alueen köyhyys ja tarkastellaan, minkälainen yhteys maakunnittain lasketulla köyhyydellä on pitkäaikaissairastavuuteen. Estimointitulokset esitetään liitteessä 8. Kuten edellä pitkäaikaissairauksien kohdalla myös yleisen terveydentilan ollessa selitettävänä muuttujana sekä päälukumitta että köyhyyskuilu saavat ennako-odotusten vastaisesti negatiivisen kertoimen. Kertoimet jäävät kuitenkin vaille tilastollista merkitsevyyttä, joten terveydentilalla ja maakunnan köyhyydellä ei tulosten mukaan ole yhteyttä. Toisaalta hypoteesia köyhyyden terveyttä huonontavasta vaikutuksesta ei voida kuitenkaan luotettavasti kumota, koska köyhyysmittojen negatiivinen parametriestimaatti jää malleissa vaille tilastollista merkitsevyyttä. Köyhyyden huomioiminen ei kuitenkaan muuta edellä esitettyjä tuloksia tuloerojen ja yleisen terveydentilan välisestä yhteydestä. Tuloroilla ja yleisellä terveydentilalla on edelleen tilastollisesti merkitsevä negatiivinen yhteys vuonna 2003. Vuonna 2004 yhteyttä ei esiinny.

## 7 YHTEENVETO JA JOHTOPÄÄTÖKSET

Tämän tutkielman tavoitteena oli selvittää, minkälainen tuloerojen ja terveyden välinen yhteys on Suomessa. Tutkimusongelma tiivistettiin muutamaan aiheen kannalta keskeiseen kysymykseen ja kirjallisuudesta johdettuun hypoteesiin. Tutkielmassa selvitettiin, minkälainen yhteys maakunnittain lasketuilla tuloroilla oli yksilöiden pitkäaikaissairastavuuteen ja itse arvioituun yleiseen terveydentilaan Suomessa vuosina 2003 ja 2004. Tutkielman ensimmäinen hypoteesi johdettiin tulojen eriarvoisuuden hypoteesin vahvasta versiosta, jonka mukaan tuloterot ovat haitallisia yksilöiden terveydelle. Lisäksi



tutkimuksessa selvitettiin, oliko maakunnan tuloerojen yhteys yksilöiden pitkäaikaissairastavuuteen ja yleiseen terveydentilaan erilainen eri tulotasoilla olevien henkilöihin välillä vuosina 2003 ja 2004. Tutkielman toinen hypoteesi johdettiin tulojen eriarvoisuuden hypoteesin heikosta versiosta, jonka mukaan tuloerot ovat haitallisimpia tulojen suhteen huono-osaisten terveydelle.

Tutkielmassa aiheen teoreettinen tarkastelu jakautui kolmeen osaan. Ensimmäisessä osassa (luku 2) selvitettiin tuloerojen ja köyhyyden mittaamisen teoriaa, ja luotiin katsaus tuloeroissa, köyhyydessä ja kansanterveydessä Suomessa viime vuosina tapahtuneisiin muutoksiin. Toinen osa (luku 3) käsitteli tuloerojen ja terveyden välistä yhteyttä kirjallisuudessa esitettyjen hypoteesien ja vaikutusmekanismien näkökulmasta. Keskeisenä tarkastelukohteena oli tutkielman empiirisessä osassa testattava tulojen eriarvoisuuden hypoteesi. Kolmannessa osassa (luku 4) käytiin läpi viime vuosina julkaistuja pohjoismaalaisia tutkimuksia tuloerojen yhteydestä terveyteen ja kuolleisuuteen.

Tutkielman empiirisen osan (luku 6) tavoitteena oli testata ekonometrisen analyysin avulla tulojen eriarvoisuuden hypoteesin vahvaa ja heikkoa versiota sekä selvittää, muuttuuko tulos tuloerojen ja terveyden välisestä yhteydestä, mikäli samanaikaisesti otetaan huomioon alueellinen köyhyys. Aineistona käytettiin Tilastokeskuksen vuosien 2003 ja 2004 tulonjakotilastoja, joihin yhdistettiin Eurostatin tulo- ja elinolotutkimusta (EU-SILC) varten kerätystä SILC-aineistosta yli 15-vuotiaiden suomalaisten pitkäaikaissairastavuutta ja yleistä terveydentilaa koskevat tiedot. Tulonjakotilasto on edustava otos Suomen väestöstä, ja siitä saadaan kattavat tiedot suomalaisten kotitalouksien tuloista ja niiden kehityksestä, sekä voidaan tehdä luotettavia laskelmia tuloeroista ja niiden kehityksestä eri alueluokituksilla. Vuodesta 2003 lähtien tulonjakotilaston tietojen keruun yhteydessä on kerätty tietoja Eurostatin tulo- ja elinolotutkimusta varten. Tässä tutkimuksessa hyödynnettiin ensimmäistä kertaa Suomessa SILC-aineistosta saatavia tietoja tuloerojen ja terveyden välisen yhteyden tutkimisessa. Aineiston ekonometrisen analyysin suoritettiin probit-mallia ja järjestettyä probit-mallia käyttäen.

Tutkielman empiirisen analyysin keskeinen havainto oli, että maakunnan tuloeroilla ja pitkäaikaissairastavuudella ei ollut yhteyttä Suomessa vuosina 2003 ja 2004, eikä tuloeroilla havaittu suurempaa pitkäaikaissairastavuutta lisäävää vaikutusta tulojen suhteen huono-osaisten kohdalla, kun tuloeroja mitattiin Gini-kertoimella. Tulokset eivät näiltä osin tukeneet

tulojen eriarvoisuuden hypoteesin vahvaa eivätkä heikkoa versiota (hypoteesit 1 ja 2). Tulokset maakunnan tuloerojen ja pitkäaikaissairastavuuden välisestä yhteydestä säilyivät pääpiirteissään muuttumattomina, vaikka samanaikaisesti otettiin huomioon alueen köyhyys. Eräs selitys hypoteesin vastaiselle tulokselle voi löytyä psykososiaalisen selitysmallin kautta. Objektiivisena terveyden mittarina pitkäaikaissairastavuus ei välttämättä heijasta riittävästi esimerkiksi yksilöiden tuntemuksia tuloerojen kautta mahdollisesti heikentyvästä sosiaalisesta pääomasta, ja sosiaalisten hierarkioiden aiheuttamasta lisääntyvästä stressistä. Toisaalta on mahdollista, että maakunnittain lasketut tuloerot eivät kuvaa riittävästi yksilöiden välistä taloudellista eriarvoisuutta, eikä tuloerojen ja terveyden välistä yhteyttä tästä syystä esiinny.

Seutukuntien välillä tuloerot vaihtelivat aineistossa maakuntia enemmän, ja seutukuntakohtaisilla tuloeroilla ja pitkäaikaissairastavuudella saatiin maakuntatarkastelusta poiketen tilastollisesti merkitsevä positiivinen yhteys vuonna 2003. Tutkielman tulos ei näiltä osin vastaisi esitettyä väitettä siitä, että maantieteellisesti pienille alueille lasketut tuloerot eivät heijastaisi riittävästi yhteiskunnallista eriarvoisuutta, eikä tuloerojen yhteyttä terveyteen tästä syystä esiintyisi (ks. Subramanian & Kawachi 2004, 82). Tulokinnassa on kuitenkin oltava varovainen, sillä seutukunnan tuloerojen ja pitkäaikaissairastavuuden välillä ei esiintynyt yhteyttä vuonna 2004.

Tutkielman empiirisen osan toinen keskeinen havainto oli, että maakunnan tuloeroilla ja yleisellä terveydentilalla oli negatiivinen yhteys vuonna 2003, kun tuloeroja mitattiin Gini-kertoimella. Tuloerojen ja yleisen terveydentilan välillä oli negatiivinen yhteys riippumatta estimoinnissa käytetystä eriarvoisuusmitasta. Tulos säilyi ennallaan, vaikka selitettävää muuttuja korvattiin yleiseen terveydentilaan yhdistettävällä terveyteen liittyvää elämänlaatua mittaavalla 15D-terveysmittarilla. Tulokset tuloerojen ja yleisen terveydentilan välisestä yhteydestä eivät myöskään muuttuneet, vaikka samanaikaisesti huomioitiin maakunnan köyhyys. Vuoden 2003 osalta tulokset tukivat tulojen eriarvoisuuden hypoteesin vahvaa versiota (hypoteesi 1). Eräs selitys pitkäaikaissairastavuuden osalta saaduista tuloksista poikkeavalle havainnolle voi löytyä psykososiaalisen selitysmallin kautta. Pitkäaikaissairastavuutta subjektiivisempaan terveyden indikaattorina yleinen terveydentila voi heijastaa yksilöiden tuntemuksia terveydestään paremmin kuin objektiivisesti todettuihin sairauksiin perustuvat terveysmittarit. Subjektiiviseen arvioon perustuva terveyden indikaattori voi reagoida herkemmin esimerkiksi tuloerojen synnyttämien sosiaalisten hierarkioiden aiheuttamiin tuntemuksiin ja ihmisten keskinäisen vertailun synnyttämään



stressiin. Tutkielman tulokset osoittivat kuitenkin, että tuloerojen ja yleisen terveydentilan välinen yhteys ei vaihdellut yksilöiden suhteellisen tulotason mukaan. Tulokset eivät tukeneet tulojen eriarvoisuuden hypoteesin heikkoa versiota (hypoteesi 2), kun selitettävänä muuttujana oli yleinen terveydentila. Tulokset vastasivat tältä osin pitkäaikaissairastavuuden ja tuloerojen välisestä yhteydestä tehtyjä havaintoja.

Yleisen terveydentilan osalta seutukuntakohtainen tarkastelussa tulokset eivät tukeneet tulojen eriarvoisuuden hypoteesia, mutta seutukunnan tuloeroille ja terveyteen liittyvää elämän laatua kuvaavalle 15D-terveysmittarille saatiin tutkielmassa tilastollisesti merkitsevä käänteinen yhteys vuonna 2003. Tuloksen mukaan seutukunnan tuloerojen kasvu johti pienempään 15D-mittarin arvoon, ja sitä kautta heikentyneeseen terveyteen liittyvään elämänlaatuun. Eräs selitys yleisestä terveydentilasta poikkeavalle tulokselle voi löytyä muuttujien erilaisista määritelmistä. 15D-mittari on nollan ja yhden välillä arvoja saava jatkuva muuttuja, joka perustuu 15 etukäteen määritellyyn terveyden osa-alueeseen. Mittari voi siten kuvata koettua terveydentilaa täsmällisemmin kuin yksilön oma arviointi koetusta yleisestä terveydentilasta sellaisenaan, ja reagoida terveydentilassa tapahtuviin muutoksiin herkemmin.

Muiden pitkäaikaissairastavuutta ja yleistä terveydentilaa kuvaavien kontrollimuuttujien osalta tulokset vastasivat pitkälti ennakko-odotuksia. Tulosten mukaan sekä maakunnan että seutukunnan keskimääräisellä tulolla oli positiivinen yhteys yleiseen terveydentilaan: terveydentila oli parempi niillä alueilla, joissa keskimääräinen käytettävissä oleva tulo oli korkeampi. Yksilön käytettävissä olevilla tuloilla ja koulutuksella havaittiin odotetusti terveyttä suojaava vaikutus. Tulojen kasvu ja kouluttautuneisuus pienensivät sekä pitkäaikaissairastavuuden todennäköisyyttä että paransivat itse arvioitua yleistä terveydentilaa. Sosioekonomisen aseman osalta tuloksista havaittiin, että pitkäaikaistyöttömät ja eläkeläiset olivat sairaampia ja kokivat yleisen terveydentilansa muita sosioekonomisia ryhmiä huonommaksi. Yleinen terveydentila oli ylemmillä ja alemmilla toimihenkilöillä muita ryhmiä parempi. Odotusten mukaisesti myös ihmissuhteilla ja sosiaalisella verkostolla oli terveyttä suojaava vaikutus: naimisissa olevat ja suurissa perheissä asuvat henkilöt olivat muita terveempiä. Toisaalta pitkäaikaissairastavuuden todennäköisyys kasvoi ja yleinen terveydentila heikkeni odotetusti iän myötä. Tuloksista havaittiin myös, että naisilla oli miehiä enemmän pitkäaikaissairauksia, mutta se ei näyttänyt vaikuttavan heidän itse arvioimaansa yleiseen terveydentilaan, sillä naiset raportoivat yleisen terveydentilansa miehiä paremmaksi.

Käsillä olevassa tutkielmassa tuloerojen hypoteesia tukevaa näyttöä saatiin vuoden 2003 osalta, kun tuloeroja mitattiin Gini-kertoimella sekä maa- että seutukunnittain. Myös Norjassa tuloeroilla on havaittu yhteys terveyteen, kun sitä mitattiin kuolleisuudella (ks. Dahl ym. 2006). Useissa pohjoismaisissa tutkimuksissa tuloeroilla ja terveydellä ei ole kuitenkaan osoitettu olevan yhteyttä (ks. luvut 4.2 ja 4.3), ja on esitetty, että yhteys olisi havaittavissa ainoastaan maissa, joissa eriarvoisuus on suurta (ks. Subramanian & Kawachi 2004). Tätä näkemystä ovat tukeneet myös Suomessa aiemmin julkaistut tutkimukset, joissa tuloerojen nousun ei ole todettu vaikuttavan haitallisesti terveyteen (ks. Blomgren ym. 2004, Martikainen ym. 2004, Böckerman ym. 2007). Käsillä olevaan tutkielmaan nähden poikkeavat tulokset voivat olla seurausta esimerkiksi tutkimusten erilaisista terveyden indikaattoreista. Kuten Böckerman ym. (2007, 3) huomauttavat, täsmällistä tietoa siitä, mihin terveyden indikaattoreihin tuloerot vaikuttavat, ei ole. Tässä tutkielmassa saatiin joitakin viitteitä siitä, että maakunnan tuloerot vaikuttivat haitallisemmin subjektiivisesti arvioituun yleiseen terveydentilaan kuin objektiivisemmin todennettuihin sairauksiin tai vammoihin perustuvaan terveyden indikaattoriin vuonna 2003. Böckerman ym. (2007) havaitsivat, että naisten kohdalla tuloerojen kasvulla oli käänteinen yhteys itse arvioituun hyvään fyysiseen terveyteen Suomessa. Terveyttä kuvaavan mittarin ja sukupuolen vaikutusta tuloerojen ja terveyden väliseen yhteyteen olisikin syytä tutkia lisää. Käsillä olevan tutkielman hypoteesia osittain tukeneet havainnot antavat Böckermanin ym. (2007) naisia koskeneiden tulosten ohella viitteitä myös siitä, että tuloerojen yhteyttä terveyteen tasaisen tulonjaon maissa, kuten Suomessa ja muissa Pohjoismaissa olisi selvitettävä tarkemmin.

Vaikka tutkielmassa tuloerojen hypoteesia tukevaa näyttöä saatiin vuoden 2003 osalta, ei pitkäaikaissairastavuudella ja yleisellä terveydentilalla esiintynyt tilastollisesti merkitsevää yhteyttä tuloeroihin vuotta myöhemmin. Tuloerojen ja terveyden yhteyttä koskevien johtopäätösten teossa pelkästään yhden vuoden poikkileikkausaineiston perusteella on oltava varovainen. Blomgren ym. (2004), Martikainen ym. (2004) ja Böckerman ym. (2007) hyödynsivät tutkimuksissaan joko usean vuoden paneeliaineistoa tai yhdistettyä aineistoa. Paneeliaineistolla tehty tarkastelu ei sisältynyt käsillä olevan tutkielmaan, mutta tutkielman aineistolla olisi mahdollista analysoida kahden vuoden paneelia. Paneeliaineiston avulla tehtävät lisätutkimukset tuloerojen yhteydestä esimerkiksi subjektiivisesti arvioituun terveydentilaan voisivat olla hyödyllisiä, mikäli halutaan tarkastella ajan myötä tapahtuvia muutoksia ja saada tietoa mahdollisista kausaalisuusvaikutuksista tuloerojen ja terveyden välillä. Käsillä olevan tutkielman eräs rajoitus on, että sen empiiriset testit kuvaavat



ensisijaisesti tuloerojen ja terveyden välistä korrelaatio, eikä niiden perusteella voida tehdä päätelmiä kausaalisuussuhteista.

Tutkielman tuloksia tarkasteltaessa on myös otettava huomioon, että maakunta ja seutukunta eivät välttämättä ole oikeita viiteryhmiä, joihin yksilöiden vertailu kohdistuu ja, jonka suhteen tuloeroja tulisi tarkastella. Teorian mukaan yksilön viiteryhmä voi olla yhtä hyvin mikä tahansa muu ryhmä, johon yksilö kokee kuuluvansa (ks. luku 3.2). Tuloerot voitaisiin siten laskea esimerkiksi kohorteittain (vrt. Deaton 1999b), ja selvittää, onko tuloeroilla ja terveydellä tällöin yhteyttä. Lisäksi suomalaisissa tutkimuksissa on tuloerojen laskemisessa käytetty aiemmin pelkästään Gini-kerrointa, joka on vain yksi, joskin yleisimmin käytetty tuloerojen mittari. Olisi tärkeää selvittää tarkemmin, minkälainen tuloerojen ja terveyden yhteys on, kun tuloerojen laskemisessa käytetään muita eriarvoisuusmittoja tai ekvivalenssiskaaloja. Käsillä olevassa tutkielmassa havaittiin, että tulokset tuloerojen ja pitkäaikaissairastavuuden välisestä yhteydestä muuttuivat, kun Gini-kerroin korvattiin Theilin indeksillä ja, kun per capita -skaalan sijaan käytettiin muita ekvivalenssiskaaloja. Tutkimusongelmaa olisi hyödyllistä selvittää lisää myös niin sanottujen monitasomallien avulla, sillä niiden avulla olisi mahdollista tutkia esimerkiksi sitä, vaihtelee tuloerojen ja terveyden välinen yhteys eri alueiden välillä.

Käsillä olevan tutkielman tulokset antoivat joitakin viitteitä tuloerojen kasvun terveyttä heikentävästä vaikutuksesta. Tuloksien tulkintaan on kuitenkin suhtauduttava kriittisesti, sillä tuloerojen ja terveyden välisen yhteyden tutkimiseen liittyy monia rajoitteita ja vielä ratkaisemattomia kysymyksiä. Suomessa tuloerojen ja terveyden välistä yhteyttä on tutkittu varsin vähän siitä huolimatta, että tuloerot ovat kasvaneet merkittävästi viime vuosina ja väestöryhmien väliset terveyserot ovat huomattavan suuret. Aiheesta tarvittaisiin lisää tietoa, jotta tutkimustuloksista voidaan tehdä päteviä johtopäätöksiä. Tutkimus tuloerojen ja terveyden välisestä yhteydestä Suomessa on epäilemättä taipaleensa alkuvaiheessa, ja aiheesta tehdyille tutkimuksille riittäisi tilaa.

## LÄHTEET

Atkinson, A. B. 1970. On the measurement of inequality. *Journal of Economic Theory* 2. 244–263.

Atkinson, A. B. 1998. *Poverty in Europe*. Yrjö Johansson Lectures. Blackwell, Oxford.

Blomgren, Jenni & Martikainen, Pekka & Mäkelä, Pia & Valkonen, Tapani 2004. The effects of regional characteristics on alcohol-related mortality – a register-based multilevel analysis of 1,1 million men. *Social Science & Medicine* 58. 2523–2535.

Blomgren, Jenni 2005. *Huono-osaisuus Suomen kaupunkiseutukunnissa: alue-erot ja sosiaalisen ympäristön vaikutukset 1990-luvulla*. Helsingin yliopiston sosiologian laitoksen tutkimuksia nro 246. Yliopistopaino, Helsinki.

Böckerman, Petri & Johansson, Edvard & Helakorpi, Satu & Uutela, Antti 2007. *Economic inequality and health: looking beyond aggregate indicators*. Keskustelualoitteita 1104. Elinkeinoelämän tutkimuslaitos, Helsinki.

Cowell, Frank A. 1995. *Measuring Inequality*. 2<sup>nd</sup> edition. Prentice Hall/Harvester Wheatsheaf, Lontoo.

Cutler, David M. & Deaton, Angus S. & Lleras-Muney, Adriana 2006. *The Determinants of Mortality*. Working Paper 11963. National Bureau of Economic Research, Cambridge.

Dahl, Espen & Elstad, Jon Ivar & Hofoss, Dag & Martin-Mollard, Melissa 2006. For whom is income inequality most harmful? A multi-level analysis of income inequality and mortality in Norway. *Social Science & Medicine* 63. 2562–2574.

Deaton, Angus 1999a. *Inequalities in health and inequalities in income*. Working Paper 7141. National Bureau of Economic Research, Cambridge.

Deaton, Angus 1999b. *Mortality, education, income and inequality among American cohorts*. Working Paper 7140. National Bureau of Economic Research, Cambridge.



Deaton, Angus 2001. *Relative deprivation, inequality, and mortality*. Working Paper 8099. National Bureau of Economic Research, Cambridge.

Deaton, Angus 2002. Policy Implications Of The Gradient Of Health And Wealth. *Health Affairs* 21. 13–30.

Deaton, Angus 2003. Health, Inequality, and Economic Development. *Journal of Economic Literature* 41. 113–158.

Deaton, Angus 2004. *Health in an age of globalization*. Working Paper 10669. National Bureau of Economic Research, Cambridge.

Duclos, Jean-Yves & Araar, Abdelkrim 2006. *Poverty and Equity: Measurement, Policy and Estimation with DAD*. Springer, New York.

Euroopan komissio 2007. *Lot 4: Methodological Paper N°1 Measuring health*. Henkilökohtainen tiedoksianto Tilastokeskuksen erikoistutkija Hannele Saulilta 26.3.2007.

Foster, James & Greer, Joel & Thorbecke, Erik 1984. A Class of Decomposable Poverty Measures. *Econometrica* 52. 761–766.

Foster, James E. 1998. Absolute versus Relative Poverty. *The American Economic Review* 88. 335–341.

Gerdtham, Ulf-G. & Johannesson, Magnus 2004. Absolute Income, Relative Income, Income Inequality, and Mortality. *Journal of Human Resources* 39. 229–247.

Gravelle, Hugh 1999. How much of the relation between population mortality and unequal distribution of income is a statistical artefact? Teoksessa Kawachi, Ichiro & Kennedy, Bruce P. & Wilkinson, Richard G. (toim.) *The society and population health reader. Income inequality and health*. The New Press, New York. 99–104.

Greene, William H. 2003. *Econometric analysis*. 5<sup>th</sup> edition. Prentice Hall, Upper Saddle River.

Hagfors, Robert 1993. Desiiliryhmittelystä tulonjakotutkimuksessa. *Kansantaloudellinen aikakauskirja* 89. 414–416.

Helakorpi, Satu & Laitalainen, Elina & Absetz, Pilvikki & Torppa, Jorma & Uutela, Antti & Puska, Pekka 2007. *Aikuisväestön terveyskäyttäytyminen ja terveys maakunnissa 1978–2005*. Kansanterveyslaitoksen julkaisu B 15/2007. Edita Prima Oy, Helsinki.

Henriksson, Göran & Allebeck, Peter & Ringbäck, Weitoft & Thelle, Dag 2006. Income distribution and mortality: Implications from a comparison of individual-level analysis and multilevel analysis with Swedish data. *Scandinavian Journal of Public Health* 34. 287–294.

Hildebrand, Vincent & Van Kerm, Philippe 2005. *Income Inequality and Self-Rated Health Status: Evidence from the European Community Household Panel*. SEDAP Research Paper No. 127. McMaster University, Hamilton.

Jenkins, Stephen P. 1999. Analysis of income distributions. *Stata Technical Bulletin* 48. 4–18.

Kansaneläkelaitos 2007. *Kelan sairastavuusindeksi kertoo kuntien terveystilanteen*. <http://www.kela.fi/in/internet/suomi.nsf/NET/080904082237PN?openDocument>, 4.9.2007.

Kautto, Mikko & Fritzell, Johan & Hvinden, Bjørn & Kvist, Jon & Uusitalo, Hannu 2004. Conclusion: Nordic welfare states in the European context. Teoksessa Kautto, Mikko & Fritzell, Johan & Hvinden, Bjørn & Kvist, Jon & Uusitalo, Hannu (toim.) *Nordic welfare states in the European context*. Routledge, New York. 262–272.

Kawachi, Ichiro & Kennedy, Bruce P. 1999. Income Inequality and Health: Pathways and Mechanisms. *Health Services Research* 34. 215–227.

Lambert, Peter J. 1993. *The Distribution and Redistribution of Income: A Mathematical Analysis*. Manchester University Press, Manchester.

Li, Hongbin & Zhu, Yi 2006. *Income, Income Inequality and Health. Evidence from China*. Discussion Paper No. 2006/07. World Institute for Development Economics Research, Helsinki.



Loikkanen, Heikki A. & Laakso, Seppo & Sullström, Risto 1997. *Alueellisista tuloeroista Suomessa. Osa I*. VATT-Keskustelualoitteita 152. J-Paino Oy, Helsinki.

Loikkanen, Heikki A. & Riihelä, Marja & Sullström, Risto 2005. Regional Income Convergence and Inequality in Boom and Bust: Results from Micro Data in Finland 1971-2000. Teoksessa Felsenstein, D. & Portnov, B. A. (toim.) *Regional Disparities in Small Countries*. Springer, Berlin/Heidelberg. 109–127.

Lynch, John & Davey Smith, George & Harper, Sam & Hillemeier, Marianne & Ross, Nancy & Kaplan, George A. & Wolfson, Michael 2004. Is Income Inequality a Determinant of Population Health? Part 1. A Systematic Review. *The Milbank Quarterly Journal* 82. 5–99.

Lyngstadt, Jan & Keilman, Nico & Bojer, Hilde & Thomsen, Ib 1997. Poverty and Economic Inequality: Concepts, Measures, and Methods. Teoksessa Keilman, Nico & Lyngstadt, Jan & Bojer, Hilde & Thomsen, Ib (toim.) *Poverty and Economic Inequality in Industrialized Western Societies*. Scandinavian University Press, Oslo.

Macinko, James A. & Shi, Leiyu & Starfield, Barbara & Wulu, John T. Jr. 2003. Income Inequality and Health: A Critical Review of the Literature. *Medical Care Research and Review* 60. 407–452.

Maddala, G. S. 1983. *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press, Cambridge.

Manderbacka, Kristiina 2005. Koettu terveys ja tiedossa oleva sairastavuus. Teoksessa Aromaa, Arpo & Huttunen, Jussi & Koskinen, Seppo & Teperi, Juha (toim.) *Suomalaisten terveys*. Kustannus Oy Duodecim, Helsinki. 130–133.

Martelin, Tuija & Koskinen, Seppo & Lahelma, Eero 2005. Väestöryhmien väliset terveyserot. Teoksessa Aromaa, Arpo & Huttunen, Jussi & Koskinen, Seppo & Teperi, Juha (toim.) *Suomalaisten terveys*. Kustannus Oy Duodecim, Helsinki. 266–276.

Martelin, Tuija & Koskinen, Seppo & Valkonen, Tapani 2005. Kuolleisuus. Teoksessa Aromaa, Arpo & Huttunen, Jussi & Koskinen, Seppo & Teperi, Juha (toim.) *Suomalaisten terveys*. Kustannus Oy Duodecim, Helsinki. 117–129.

Martelin, Tuija & Koskinen, Seppo & Valkonen, Tapani 2006. Mortality. Teoksessa Koskinen, Seppo & Aromaa, Arpo & Huttunen, Jussi & Teperi, Jussi (toim.) *Health in Finland*. Vammalan Kirjapaino Oy, Vammala. 48–51.

Martikainen, Pekka & Mäki, Netta & Blomgren, Jenni 2004. The Effects of Area and Individual Social Characteristics on Suicide Risk: A Multilevel Study of Relative Contribution and Effect Modification. *European Journal of Population* 20. 323–350.

Mellor, Jennifer M. & Milyo, Jeffrey 2002. Income Inequality and Health Status in the United States. Evidence from the Current Population Survey. *The Journal of Human Resources* 37. 510–539.

Miller, Douglas L. & Paxson, Christina 2001. *Relative income, race and mortality*. Working Paper 13. Princeton University, Princeton.

Niemi, Erkki 2004. Kaupunkien rajat. Laitakaupunki, esikaupunki, nukkumalähiö. *Hyvinvointikatsaus* 3/2004. 4–13.

Osler, Merete & Prescott, Eva & Grønbaek, Morten & Christensen, Ulla & Due, Pernille & Engholm, Gerda 2002. Income inequality, individual income, and mortality in Danish adults: analysis of pooled data from two cohort studies. *British Medical Journal* 324. 13–16.

Preston, Samuel H. 1975. The changing relation between mortality and level of development. *Population Studies* 29. 231–248.

*Pääministeri Matti Vanhasen II hallituksen ohjelma 2007*. Edita Prima Oy, Helsinki.

Rahkonen, Ossi & Talala, Kirsi & Laaksonen, Mikko & Lahelma, Eero & Prättälä, Ritva & Uutela, Antti 2004. Suomalaisten koettu terveys parantunut, terveyden koulutuserot säilyneet 1979–2002. *Suomen Lääkärilehti* 20. 2159–2163.



Riihelä, Marja & Sullström, Risto 2001. *Tuloerot ja eriarvoisuus suuralueilla pitkällä aikavälillä 1971–1998 ja erityisesti 1990-luvulla*. VATT-tutkimuksia 80. Oy Nord Print Ab, Helsinki.

Riihelä, Marja & Sullström, Risto & Tuomala, Matti 2005. *Trends in Top Income Shares in Finland*. VATT-keskustelualoitteita 371. Oy Nord Print Ab, Helsinki.

Riihelä, Marja & Sullström, Risto & Tuomala, Matti 2007. *Economic poverty in Finland 1971–2004*. VATT-Keskustelualoitteita 418. Oy Nord Print Ab, Helsinki.

Rodgers, G. B. 1979. Income and inequality as determinants of mortality: an international cross-section analysis. *Population Studies* 33. 343–351.

Rostgaard, Tine & Lehto, Juhani 2004. Health and social care systems: How different is the Nordic model? Teoksessa Kautto, Mikko & Fritzell, Johan & Hvinden, Bjørn & Kvist, Jon & Uusitalo, Hannu (toim.) *Nordic welfare states in the European context*. Routledge, New York. 137–167.

Ruotsalainen, Pekka 2005. Tuloerot kasvussa ympäri Suomen. *Hyvinvointikatsaus* 1/2005. 20–25.

Schmidt, Helena 2007. Terveystilan eriytyminen hierarkkisten yhteiskuntien välillä ja sisällä. (kuvio) Teoksessa Pasternack, Amos & Pelkonen, Risto & Eskola, Kari & Haahtela, Tari (toim.) *Lääkäriksi*. Kustannus Oy Duodecim, Helsinki.

Shi, Leiyu & Starfield, Barbara 2000. Primary care, income inequalities, and self-rated health in the United States: A mixed-level analysis. *International Journal of Health Services* 30. 541–555.

Sintonen, Harri 2001. The 15D instrument of health-related quality of life: properties and applications. *Annals of Medicine* 33. 328–336.

Skalli, Ali & Johansson, Edvard & Theodossiou, Ioannis 2006. *Are healthier wealthier or the wealthier healthier*. Yliopistopaino, Helsinki.

Sosiaali- ja terveysministeriö 2001. *Valtioneuvoston periaatepäätös Terveys 2015-kansanterveysohjelmasta*. Sosiaali- ja terveysministeriön julkaisuja 2001:4. Sosiaali- ja terveysministeriö, Helsinki.

Starfield, B. & Shi, L. 1999. Determinants of health: Testing of a conceptual model. *Annals of the New York Academy of Sciences* 896. 344–346.

Subramanian, S. V. & Kawachi, Ichiro 2004. Income Inequality and Health: What Have We Learned So Far? *Epidemiologic Reviews* 26. 78–91.

Suoniemi, Ilpo 2005. Yleiskuva köyhyyden kehityksestä Suomessa vuosina 1990–2002. *Talous & Yhteiskunta* 2. 23–29.

Talousneuvosto 2001. *Työmarkkinoilta syrjäytyminen, tulonjako ja köyhyys*. Työryhmäraportti. Valtioneuvoston kanslian julkaisusarja 2001/13.

Theil, H. 1967. *Economics and Information Theory*. North Holland, Amsterdam.

Tilastokeskus 2004. *Kunnat ja kuntapohjaiset aluejaot 2004*. Käsikirjoja 28. Yliopistopaino, Helsinki.

Tilastokeskus 2006a. *Kotitalouksien tulot ja tuloerot maakunnissa 1995–2004*. [http://www.stat.fi/til/tjkt/2004/tjkt\\_2004\\_2006-02-28\\_kat\\_001.html](http://www.stat.fi/til/tjkt/2004/tjkt_2004_2006-02-28_kat_001.html), 20.9.2007.

Tilastokeskus 2006b. *Kotitalouksien väliset tuloerot Gini-kertoimella mitattuna maakunnittain 1995–2004*. [http://www.stat.fi/til/tjkt/2004/tjkt\\_2004\\_2006-02-28\\_tau\\_003.html](http://www.stat.fi/til/tjkt/2004/tjkt_2004_2006-02-28_tau_003.html), 20.9.2007.

Tilastokeskus 2006c. *Tulonjakotilasto 2004*. Multiprint Oy, Helsinki.

Tilastokeskus 2007. *SILC aineisto vuosilta 2003 ja 2004*.

Tremblay, M. 2005. *Canadian Health Measures Survey (CHMS): An introduction and overview*. Seminaariesitys Kansanterveyslaitoksella. Helsinki, Suomi.



Vuorinen, Heikki S. 2007. Taikuudesta tieteeeksi. Teoksessa Pasternack, Amos & Pelkonen, Risto & Eskola, Kari & Haahtela, Tari (toim.) *Lääkäriksi*. Kustannus Oy Duodecim, Helsinki. 67–89.

Wagstaff, Adam & van Doorslaer, Eddy 2000. Income Inequality and Health: What Does the Literature Tell Us? *Annual Review of Public Health* 21. 543–567.

Wilkinson, Richard G. 1997. Health inequalities: relative or absolute material standards? *British Medical Journal* 314, 591–595.

Wilkinson, Richard G. & Pickett, Kate E. 2006. Income inequality and population health: A review and explanation of the evidence. *Social Science & Medicine* 62. 1768–1784.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge.

World Health Organisation 1948. *Constitution of the World Health Organisation*. WHO, Geneva.

World Health Organisation 2004. *Commission on Social Determinants of Health: Note by the Secretariat*. Document number: EB115/35. World Health Organisation, Geneva.

Zimmer, Z. & Natividad, J. & Lin, H. S. & Chayovan, N. 2000. A Cross-national Examination of the Determinants of Self-assessed Health. *Journal of Health and Social Behavior* 41. 465–481.

## LIITE 1. Suomen maakunnat ja seutukunnat

### Uusimaa

Helsinki  
Lohja  
Tammisaari

### Itä-Uusimaa

Porvoo  
Loviisa

### Varsinais-Suomi

Turunmaa  
Salo  
Turku  
Vakka-Suomi  
Loimaa

### Satakunta

Rauma  
Pori  
Pohjois-Satakunta

### Kanta-Häme

Hämeenlinna  
Riihimäki  
Forssa

### Pirkanmaa

Luoteis-Pirkanmaa  
Kaakkois-Pirkanmaa  
Etelä-Pirkanmaa  
Tampere  
Lounais-Pirkanmaa  
Ylä-Pirkanmaa

### Päijät-Häme

Lahti  
Heinola

### Kymenlaakso

Kouvola  
Kotka-Hamina

### Etelä-Karjala

Lappeenranta  
Länsi-Saimaa  
Imatra  
Kärkikunnat

### Etelä-Savo

Mikkeli  
Juva  
Savonlinna  
Pieksämäki

### Pohjois-Savo

Ylä-Savo  
Kuopio  
Koillis-Savo  
Varkaus  
Sisä-Savo

### Pohjois-Karjala

Outokumpu  
Joensuu  
Ilomantsi  
Keski-Karjala  
Pielisen Karjala

### Keski-Suomi

Jyväskylä  
Joutsa  
Keuruu  
Jämsä  
Äänekoski  
Saarijärvi-Viitasaari

### Etelä-Pohjanmaa

Suupohja  
Pohjoiset seinänaapurit  
Eteläiset seinänaapurit  
Kuusiokunnat  
Härmänmaa  
Järviseu

### Pohjanmaa

Kyrönmaa  
Vaasa  
Sydösterbottens kustregion

### Keski-Pohjanmaa

Kaustinen  
Kokkola

### Pohjois-Pohjanmaa

Oulu  
Oulunkaari  
Raahe  
Siikalatva  
Nivala-Haapajärvi  
Ylivieska  
Koillismaa

### Kainuu

Kehys-Kainuu  
Kajaani

### Lappi

Rovaniemi  
Kemi-Tornio  
Tornionlaakso  
Itä-Lappi  
Tunturi-Lappi  
Pohjois-Lappi

### Ahvenanmaa

Ålands landsbygd  
Ålands landsgård

Lähde: Tilastokeskus 2004, 22–23



## LIITE 2.

Taulukko 1. Muuttujat ja niiden määritelmät

Muuttuja	Määritelmä
SELITETTÄVÄT MUUTTUJAT	
Pitkäaikaissairastavuus Yleinen terveydentila	1 jos henkilöllä on pitkäaikaissairauksia, 0 jos ei. 3 jos henkilön terveydentila on hyvä, 2 jos melko hyvä, 1 jos keskitasoinen, 0 jos melko huono tai huono.
SELITTÄVÄT MUUTTUJAT	
<i>Aluepiirteitä kuvaavat muuttujat<sup>a</sup></i>	
Keskimääräinen tulo1	Maakunnan keskimääräinen käytettävissä oleva tulo euroissa jaettuna 10 <sup>5</sup> .
Keskimääräinen tulo2	Seutukunnan keskimääräinen käytettävissä oleva tulo euroissa jaettuna 10 <sup>5</sup> .
Gini-kerroin1	Maakunnan sisäiset tuloerot Gini-kertoimella mitattuna.
Gini-kerroin2	Seutukunnan sisäiset tuloerot Gini-kertoimella mitattuna.
Atkinsonin indeksi	Maakunnan sisäiset tuloerot Atkinson-indeksillä ( $\epsilon = 1$ ) mitattuna.
Theilin indeksi	Maakunnan sisäiset tuloerot Theilin indeksillä mitattuna.
Päälukumitta	Maakunnan köyhyysaste päälukumitalla mitattuna.
Köyhyyskuilu	Köyhyysraja 60 prosenttia mediaanitulosta. Maakunnan köyhyysaste köyhyyskuilulla mitattuna. Köyhyysraja 60 prosenttia mediaanitulosta.
<i>Yksilön piirteitä kuvaavat muuttujat<sup>a</sup></i>	
Tulot	Henkilön käytettävissä olevat tulot euroina jaettuna 10 <sup>5</sup> .
Tulot <sup>2</sup>	Tulot-muuttuja korotettuna toiseen potenssiin.
R	Henkilön tulojen desiili maakunnassa.
QR	Maakuntakohtaisen Gini-kertoimen ja henkilön tulojen desiilin maakunnassa tulo.
Ikä	Ikä vuosina tarkastelukohteena olevan vuoden (2003/2004) lopussa.
Ikä <sup>2</sup>	Ikä vuosina tarkastelukohteena olevan vuoden (2003/2004) lopussa korotettuna toiseen potenssiin.
Sukupuoli	1 jos henkilö on mies, 0 jos nainen. (vertailuryhmänä naiset)
Siviilisääty	1 jos henkilö on naimisissa, 0 jos naimaton, eronnut, leski, asumus- erossa tai siviilisääty tuntematon. (vertailuryhmänä ei-naimisissa olevat henkilöt)
Koulutus	Dummy-muuttujia, jotka ilmaisevat, onko henkilö suorittanut perusasteen, keskiasteen, alimman korkea-asteen vai korkea-asteen koulutuksen. (vertailuryhmänä perusasteen koulutuksen suorittaneet)
Sosioekonominen asema	Dummy-muuttujia, jotka ilmaisevat, onko henkilö ylempi toimi- henkilö, alempi toimihenkilö, työntekijä, yrittäjä, maatalousyrittäjä, opiskelija, eläkeläinen, pitkäaikaistyötön vai muu (esim. omaa kotitaloutta hoitava henkilö). (vertailuryhmänä työntekijät)
<i>Kotitalouden piirteitä kuvaavat muuttujat</i>	
Kotitalouden koko	Kotitalouden jäsenten lukumäärä tarkasteltavana olevan vuoden (2003/2004) lopussa.

<sup>a</sup> Aluepiirteitä kuvaavien muuttujien sekä tulot- ja tulot<sup>2</sup>-muuttujan laskemisessa käytetään kolmea ekvivalenssiskaalaa, joita ovat per capita, OECD-skaala ja modifioitu OECD-skaala.

## LIITE 2. Jatkuu

Taulukko 2. Muuttujien keskiarvot, keskihajonnat, minimi ja maksimit vuonna 2003

2003				
Muuttuja	Keskiarvo	Keskihajonta	Minimi	Maksimi
<b>SELITETTÄVÄT MUUTTUJAT</b>				
Pitkäaikaissairastavuus (n=10829 )	0,393	0,488	0	1
Naiset (n=5173)	0,420	0,494	0	1
Miehet (n=5656)	0,368	0,482	0	1
Yleinen terveydentila (n=10826)	2,057	1,006	0	3
Naiset (n=5174)	2,066	1,005	0	3
Miehet (n=5652)	2,050	1,006	0	3
<b>SELITTÄVÄT MUUTTUJAT (n=10826)</b>				
<i>Aluepiirteitä kuvaavat muuttujat<sup>a</sup></i>				
Keskimääräinen tulo1	0,143	0,016	0,118	0,168
Keskimääräinen tulo2	0,143	0,019	0,094	0,196
Gini-kerroin1	0,272	0,022	0,230	0,308
Gini-kerroin2	0,267	0,036	0,161	0,413
Atkinsonin indeksi	0,120	0,021	0,085	0,159
Theilin indeksi	0,150	0,053	0,088	0,272
Päälukumitta	0,137	0,043	0,083	0,227
Köyhyyskuilu	0,030	0,011	0,012	0,061
<i>Yksilön piirteitä kuvaavat muuttujat</i>				
Tulot <sup>a</sup>	0,167	0,181	0	11,173
(Tulot <sup>2</sup> ) <sup>a</sup>	0,061	1,323	0	124,30
R	6,557	2,980	1	10
QR <sup>b</sup>	1,796	0,857	0,230	3,077
Ikä	47,061	16,591	16	96
Ikä <sup>2</sup>	2490,00	1620,92	256	9216
Sukupuoli	0,522	0,4995	0	1
Siviilisääty	0,540	0,498	0	1
<i>Koulutus:</i>				
Perusaste	0,319	0,466	0	1
Keskiaste	0,397	0,489	0	1
Alin korkea-aste	0,134	0,341	0	1
Korkea-aste	0,150	0,357	0	1
<i>Sosioekonominen asema:</i>				
Ylempi toimihenkilö	0,131	0,337	0	1
Alempi toimihenkilö	0,143	0,350	0	1
Työntekijä	0,136	0,343	0	1
Yrittäjä	0,117	0,322	0	1
Maatalousyrittäjä	0,068	0,253	0	1
Opiskelija	0,081	0,274	0	1
Eläkeläinen	0,228	0,420	0	1
Pitkäaikaistyötön	0,059	0,236	0	1
Muu	0,036	0,186	0	1
<i>Kotitalouden piirteitä kuvaavat muuttujat</i>				
Kotitalouden koko	2,579	1,405	1	16

<sup>a</sup> Ekvivalenssiskaalana per capita.

<sup>b</sup> Q = Gini-kerroin1



## LIITE 2. Jatkuu

Taulukko 3. Muuttujien keskiarvot, keskihajonnat, minimi ja maksimit vuonna 2004

2004				
Muuttuja	Keskiarvo	Keskihajonta	Minimi	Maksimi
<b>SELITETTÄVÄT MUUTTUJAT</b>				
Yleinen terveydentila (n=10904)	2,053	1,007	0	3
Naiset (n=5513)	2,064	1,005	0	3
Miehet (n=5391)	2,041	1,010	0	3
Yleinen terveydentila (n=10904)	2,053	1,007	0	3
Naiset (n=5513)	2,064	1,005	0	3
Miehet (n=5391)	2,041	1,010	0	3
<b>SELITTÄVÄT MUUTTUJAT (n=10904)</b>				
<i>Aluepiirteitä kuvaavat muuttujat<sup>a</sup></i>				
Keskimääräinen tulo1	0,151	0,018	0,123	0,181
Keskimääräinen tulo2	0,151	0,021	0,108	0,226
Gini-kerroin1	0,278	0,027	0,228	0,310
Gini-kerroin2	0,274	0,037	0,162	0,435
Atkinsonin indeksi	0,126	0,026	0,081	0,158
Theilin indeksi	0,170	0,065	0,089	0,260
Päälukumitta	0,129	0,047	0,080	0,236
Köyhyyskuilu	0,028	0,012	0,017	0,061
<i>Yksilön piirteitä kuvaavat muuttujat</i>				
Tulot <sup>a</sup>	0,171	0,212	0	14,595
(Tulot <sup>2</sup> ) <sup>a</sup>	0,074	2,199	0	213,009
R	6,401	3,015	1	10
QR <sup>b</sup>	1,793	0,887	0,228	3,010
Ikä	46,613	17,287	16	96
Ikä <sup>2</sup>	2471,54	1680,26	256	9216
Sukupuoli	0,494	0,500	0	1
Siviilisääty	0,523	0,499	0	1
<i>Koulutus:</i>				
Perusaste	0,315	0,465	0	1
Keskiaste	0,410	0,492	0	1
Alin korkea-aste	0,125	0,331	0	1
Korkea-aste	0,150	0,357	0	1
<i>Sosioekonominen asema:</i>				
Ylempi toimihenkilö	0,124	0,330	0	1
Alempi toimihenkilö	0,149	0,356	0	1
Työntekijä	0,133	0,340	0	1
Yrittäjä	0,101	0,302	0	1
Maatalousyrittäjä	0,055	0,229	0	1
Opiskelija	0,096	0,294	0	1
Eläkeläinen	0,234	0,423	0	1
Pitkäaikaistyötön	0,066	0,248	0	1
Muu	0,041	0,198	0	1
<i>Kotitalouden piirteitä kuvaavat muuttujat</i>				
Kotitalouden koko	2,581	1,416	1	16

<sup>a</sup> Ekvivalenssiskaalana per capita

<sup>b</sup> Q = Gini-kerroin1

### LIITE 3.

Taulukko 1. Pitkäaikaissairastavuuden todennäköisyyden probit-mallin kerroinestimaatit ja marginaalivaikutukset Atkinsonin ja Theilin indeksin osalta vuosina 2003 ja 2004

2003	Malli 2					
Muuttuja	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamaton keskivirhe
Atkinsonin indeksi	1,035	0,694	0,865	0,395	0,265	0,330
Havaintojen lukumäärä	10829					
Pseudo R <sup>2</sup>	0,140					
Ennustettu oikein (%)	71,18					
2004	Malli 2					
Muuttuja	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamaton keskivirhe
Atkinsonin indeksi	0,167	0,891	0,758	0,065	0,346	0,294
Havaintojen lukumäärä	10911					
Pseudo R <sup>2</sup>	0,152					
Ennustettu oikein (%)	71,40					
2003	Malli 2					
Muuttuja	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamaton keskivirhe
Theilin indeksi	0,369	0,192*	0,352	0,141	0,073	0,134
Havaintojen lukumäärä	10829					
Pseudo R <sup>2</sup>	0,140					
Ennustettu oikein (%)	71,11					
2004	Malli 2					
Muuttuja	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamaton keskivirhe
Theilin indeksi	0,171	0,441	0,383	0,066	0,171	0,149
Havaintojen lukumäärä	10911					
Pseudo R <sup>2</sup>	0,152					
Ennustettu oikein (%)	71,40					

Malleihin sisältyvät eriarvoisuusmitan lisäksi maakunnan keskimääräinen tulo –muuttuja sekä kaikki yksilö- ja kotitalouspiirteitä kuvaavat selittävät muuttujat. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon.

Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.



#### LIITE 4.

Taulukko 1. Pitkäaikaissairastavuuden todennäköisyyden probit-mallin kerroinestimaatit ja marginaalivaikutukset vuonna 2003 (tuloerot laskettu seutukunnittain)

Muuttuja	Malli 2					
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe
Gini-kerroin2	0,988	0,570*	0,464	0,377	0,218*	0,177
Keskimääräinen tulo2	-1,435	0,930	0,912	-0,547	0,356	0,348
Tulot	-0,183	0,089**	0,134	-0,070	0,034**	0,051
Tulot <sup>2</sup>	0,012	0,013	0,018	0,004	0,005***	0,007
Ikä	0,051	0,005***	0,005	0,020	0,002***	0,0020
Ikä <sup>2</sup>	0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli	-0,113	0,028***	0,027	-0,043	0,011***	0,010
Siviilisääty	-0,075	0,024***	0,032	-0,029	0,009***	0,012
Keskiaste	-0,037	0,041	0,033	-0,014	0,016	0,013
Alin korkea-aste	-0,098	0,052*	0,046	-0,037	0,019*	0,017
Korkea-aste	-0,260	0,038***	0,050	-0,096	0,013***	0,018
Ylempi toimihenkilö	0,051	0,050	0,059	0,019	0,019	0,023
Alempi toimihenkilö	-0,047	0,056	0,052	-0,018	0,021	0,020
Yrittäjä	-0,076	0,044*	0,053	-0,029	0,016*	0,020
Maatalousyrittäjä	0,128	0,061**	0,061	0,049	0,024**	0,024
Opiskelija	0,183	0,066***	0,072	0,071	0,026***	0,028
Eläkeläinen	0,703	0,049***	0,058	0,273	0,019***	0,022
Pitkäaikaistyötön	0,216	0,070***	0,063	0,084	0,028***	0,025
Muu	0,135	0,135*	0,080	0,052	0,030*	0,032
Kotitalouden koko	-0,065	-0,065***	0,012	-0,025	0,005***	0,005
Vakio	-1,779	-1,779***	0,185			
Havaintojen lkm	10829			10829		
Pseudo R <sup>2</sup>	0,140					
Ennustettu oikein (%)	71,28					

Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.

## LIITE 4. Jatkuu

Taulukko 2. Pitkäaikaissairastavuuden todennäköisyyden probit-mallin kerroinestimaatit ja marginaalivaikutukset vuonna 2004 (tuloerot laskettu seutukunnittain)

Muuttuja	Malli 2					
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe
Gini-kerroin2	0,258	0,513	0,452	0,100	0,199	0,175
Keskimääräinen tulo2	-0,719	0,918	0,833	-0,279	0,356	0,323
Tulot	-0,271	0,223	0,155	-0,105	0,087	0,060
Tulot <sup>2</sup>	0,070	0,042*	0,044	0,027	0,016	0,017
Ikä	0,045	0,006***	0,005	0,017	0,002***	0,002
Ikä <sup>2</sup>	0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli	-0,047	0,025*	0,027	-0,018	0,010*	0,011
Siviilisääty	-0,022	0,027	0,033	-0,009	0,011	0,013
Keskiaste	-0,054	0,032*	0,033	-0,021	0,012*	0,013
Alin korkea-aste	-0,126	0,043***	0,047	-0,048	0,016***	0,018
Korkea-aste	-0,216	0,048***	0,051	-0,082	0,017***	0,019
Ylempi toimihenkilö	0,039	0,047	0,060	0,015	0,018	0,023
Alempi toimihenkilö	-0,011	0,046	0,051	-0,004	0,018	0,020
Yrittäjä	-0,078	0,047*	0,055	-0,030	0,018*	0,021
Maatalousyrittäjä	0,143	0,064**	0,065	0,056	0,025**	0,026
Opiskelija	0,176	0,075**	0,066	0,069	0,030**	0,026
Eläkeläinen	0,806	0,057***	0,059	0,313	0,021***	0,022
Pitkäaikaistyötön	0,269	0,065***	0,061	0,106	0,026***	0,024
Muu	0,139	0,066**	0,076	0,054	0,026**	0,030
Kotitalouden koko	-0,047	0,013***	0,011	-0,018	0,005***	0,004
Vakio	-1,580	0,205***	0,175			
Havaintojen lkm	10911			10911		
Pseudo R <sup>2</sup>	0,152					
Ennustettu oikein (%)	71,42					

Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.



## LIITE 5.

Taulukko 1. Pitkäaikaissairastavuuden todennäköisyyden probit-mallin kerroinestimaatit vuosina 2003 ja 2004 (Malleissa on mukana alueen köyhyys.)

2003	Malli 4			Malli 5		
Muuttuja	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe
Gini-kerroin1	1,627	0,893*	1,189	1,097	0,795	1,146
Keskimääräinen tulo1	-4,636	2,219**	2,699	-2,965	1,571*	2,347
Päälukumitta	-0,776	0,725	0,718			
Köyhyyskuilu				-1,123	2,341	2,410
Havaintojen lukumäärä	10829			10829		
Pseudo R <sup>2</sup>	0,140			0,140		
Ennustettu oikein (%)	71,12			71,12		

2004	Malli 4			Malli 5		
Muuttuja	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe
Gini-kerroin1	0,261	0,919	0,925	0,366	1,082	1,021
Keskimääräinen tulo1	-0,749	1,827	2,148	-0,886	1,947	2,142
Päälukumitta	-0,215	0,599	0,591			
Köyhyyskuilu				-1,014	2,615	0,156
Havaintojen lukumäärä	10911			10911		
Pseudo R <sup>2</sup>	0,152			0,152		
Ennustettu oikein (%)	71,31			71,30		

Malleihin sisältyvät lisäksi kaikki yksilö- ja kotitalouspiirteitä kuvaavat selittävät muuttujat. Ekvivalenssisikaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon.

Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.

## LIITE 6.

Taulukko 1. Yleisen terveydentilan todennäköisyyden järjestetty probit -mallin kertoimet ja marginaalivaikutukset Atkinsonin ja Theilin indeksin osalta vuosina 2003 ja 2004

2003 Muuttuja	Malli 2			Yleinen terveydentila=0			Yleinen terveydentila=3		
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe
Atkinsonin indeksi	-1,090	0,416***	0,728	0,122	0,047***	0,082	-0,429	0,164***	0,286
Havaintojen lkm	10826								
Pseudo R <sup>2</sup>	0,106								

2004 Muuttuja	Malli 2			Yleinen terveydentila=0			Yleinen terveydentila=3		
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe
Atkinsonin indeksi	-0,166	0,780	0,634	0,019	0,090	0,072	-0,065	0,314	0,249
Havaintojen lkm	10904								
Pseudo R <sup>2</sup>	0,113								

2003 Muuttuja	Malli 2			Yleinen terveydentila=0			Yleinen terveydentila=3		
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe
Theilin indeksi	-0,548	0,089***	0,297	0,061	0,010***	0,033	-0,216	0,035***	0,117
Havaintojen lkm	10904								
Pseudo R <sup>2</sup>	0,113								

2004 Muuttuja	Malli 2			Yleinen terveydentila=0			Yleinen terveydentila=3		
	Kerroin	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe	Marginaali-vaikutus	Korjattu keskivirhe	Korjaamat. keskivirhe
Theilin indeksi	0,138	0,224	0,297	-0,016	0,026	0,033	0,054	0,088	0,117
Havaintojen lkm	10904								
Pseudo R <sup>2</sup>	0,113								

Malleihin sisältyvät eriarvoisuusmitan lisäksi maakunnan keskimääräinen tulo –muuttuja sekä kaikki yksilö- ja kotitalouspiirteitä kuvaavat selittävät muuttujat. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon. Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.



## LIITE 7.

Taulukko 1. Yleisen terveydentilan todennäköisyyden järjestetty probit -mallin kertoimet ja marginaalivaikutukset vuosina 2003 ja 2004 (tuloerot laskettu seutukunnittain)

Muuttuja	2003			2004		
	Malli 2			Malli 2		
	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe
Gini-kerroin2	-0,491	0,360	0,389	0,028	0,423	0,377
Keskimääräinen tulo2	2,990	0,730***	0,765	1,671	0,884*	0,700
Tulot	0,427	0,108***	0,117	0,490	0,148***	0,112
Tulot <sup>2</sup>	-0,037	0,017**	0,015	-0,024	0,009***	0,014
Ikä	-0,059	0,004***	0,005	-0,061	0,003***	0,004
Ikä <sup>2</sup>	0,000	0,000***	0,000	0,000	0,000***	0,000
Sukupuoli	-0,043	0,022*	0,023	-0,049	0,026*	0,023
Siviilisääty	0,082	0,023***	0,027	0,096	0,023***	0,027
Keskiaste	0,110	0,026***	0,028	0,105	0,024***	0,028
Alin korkea-aste	0,229	0,060***	0,039	0,241	0,031***	0,040
Korkea-aste	0,414	0,038***	0,043	0,356	0,036***	0,043
Ylempi toimihenkilö	0,090	0,046**	0,051	0,096	0,038**	0,052
Alempi toimihenkilö	0,111	0,038***	0,045	0,138	0,032***	0,044
Yrittäjä	-0,011	0,053	0,044	0,022	0,047	0,046
Maatalousyrittäjä	-0,206	0,053***	0,051	-0,166	0,053***	0,055
Opiskelija	-0,015	0,051	0,062	0,026	0,055	0,058
Eläkeläinen	-0,455	0,053***	0,048	-0,482	0,056***	0,048
Pitkäaikaistyötön	-0,160	0,055***	0,054	-0,249	0,051***	0,052
Muu	-0,153	0,053***	0,069	-0,179	0,072**	0,065
Kotitalouden koko	0,042	0,009***	0,010	0,039	0,011***	0,010
Havaintojen lkm	10826			10904		
Pseudo R <sup>2</sup>	0,106			0,113		

Ekvivalenssisikaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon.

Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.

## LIITE 8.

Taulukko 1. Yleisen terveydentilan todennäköisyyden järjestetty probit -mallin kerroinestimaatit vuosina 2003 ja 2004 (Malleissa on mukana alueen köyhyys.)

2003	Malli 4			Malli 5		
Muuttuja	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe
Gini-kerroin1	-1,245	0,599**	1,004	-1,230	0,546**	0,968
Keskimääräinen tulo1	4,596	1,380***	2,285	4,477	1,094***	4,477
Päälukumitta	0,251	0,406	0,604			
Köyhyyskuilu				0,865	1,288	2,032
Havaintojen lkm	10826			10826		
Pseudo R <sup>2</sup>	0,106			0,106		

2004	Malli 4			Malli 5		
Muuttuja	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe
Gini-kerroin1	-0,616	0,877	0,776	-0,375	1,032	0,857
Keskimääräinen tulo1	3,451	1,740*	1,810	2,571	1,963	1,806
Päälukumitta	0,534	0,510	0,496			
Köyhyyskuilu				1,020	2,195	1,942
Havaintojen lkm	10904			10904		
Pseudo R <sup>2</sup>	0,113			0,113		

Malleihin sisältyvät lisäksi kaikki tutkielman yksilö- ja kotitalouspiirteitä kuvaavat selittävät muuttujat. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon.

Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.



## LIITE 9.

Taulukko 1. Yleisen terveydentilan jakauma ja 15D-mittarin luokkakeskisarvot

Muuttuja	2003	2004	15D:n luokkakeskisarvo <sup>a</sup>
	Havaintojen lkm	Havaintojen lkm	
Yleinen terveydentila huono	244	266	0,703897044
Yleinen terveydentila melko huono	713	730	0,777402321
Yleinen terveydentila keskitasoinen	2310	2273	0,876937019
Yleinen terveydentila melko hyvä	2713	2796	0,927355101
Yleinen terveydentila hyvä	4846	4839	0,962036262
Yhteensä	10826	10904	

<sup>a</sup>15D-mittarin koettua terveydentilaa kuvaavat luokkakeskisarvot perustuvat Sosiaali- ja terveysalan tutkimus- ja kehittämiskeskuksen Stakesin Terveystaloustieteen keskuksen tekemiin laskelmiin Terveys 2000-tutkimuksesta, jonka terveydentilaa koskeva haastattelukysymys vastaa SILC-aineiston tiedonkeruussa käytettyä kysymystä.

Taulukko 2. Lineaarisen regressiomallin kerroinestimaatit

2003 Muuttuja	Malli 2 (maakunta)			Malli 2 (seutukunta)		
	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe
Gini-kerroin1	-0,072	0,024***	0,034			
Gini-kerroin2				-0,038	0,017**	0,018
Havaintojen lukumäärä	10826			10826		
R <sup>2</sup>	0,210			0,210		

2004 Muuttuja	Malli 2 (maakunta)			Malli 2 (seutukunta)		
	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe	Kerroin	Korjattu keskivirhe	Korjaamaton keskivirhe
Gini-kerroin1	-0,011	0,033	0,026			
Gini-kerroin2				0,001	0,021	0,018
Havaintojen lukumäärä	10904			10904		
R <sup>2</sup>	0,223			0,223		

Selitettävä muuttuja on terveyteen liittyvää elämänlaatua kuvaava 15D-mittari. Malleihin sisältyvät Gini-kertoimen lisäksi alueen keskimääräinen tulo -muuttuja sekä kaikki yksilö- ja kotitalouspiirteitä kuvaavat selittävät muuttujat. Ekvivalenssiskaalana on per capita. Taulukossa esitetään havaintojen klusteroinnin huomioivien keskivirheiden lisäksi myös korjaamattomat keskivirheet. Kertoimien tilastollisen merkitsevyyden testaus korjaamattomien keskivirheiden perusteella ei sisälly taulukkoon.

Tilastollisesti merkitsevä 10 %:n (\*), 5 %:n (\*\*) ja 1 %:n (\*\*\*) luottamustasolla.